A Novel Reading Technique for Visually Impaired Person Using Enhanced Optical Character Recognition Method

R. Rajan[1*],M. S. Geetha Devasena[2]

Corresponding Author:R. Rajan. Email: rajanmailindia@gmail.com

Address:1*Department of Computer Science and Engineering, KGISL Institute of Technology, 2Department of Computer Science and Engineering, Sri Ramakrishna Engineering College, Coimbatore,

Relevant conflicts of interest/financial disclosures: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abstract: Person with vision is able to read any document even knowing title, subsection, and paragraph as the ways. In the case of a visually impaired person, it was possible only for some other person giving reading assistants. There must be a provision for visually impaired persons who are supposed to get assistance to read content of documents or books with their title, subtitle, paragraph etc. as per their wish and themselves using some Machine learning (ML) or Deep Learning (DL) algorithms. Some work has been proposed to assist visually impaired persons. With the help of these works are able to scan documents using Object Character Recognition (OCR) technique and with the help of voice assistants, it can be listened to by them. The limitation is that the OCR technique does not identify and categorize as title, sub title and paragraph etc. The idea of this proposed work was that it is essential to identify and categorize reading image based content or document into title, sub title and paragraph within a single image. This solution would be very useful to visually impaired people to read such image content as per their choice and that they can get only what content they want to read and which title, which subtitle and paragraph and all. In this proposed work, this proposed work was coined the Enhanced OCR (EOCR) method. The experiment was conducted and accuracy of categories image content into title, subtitle, paragraph etc. for 10 in count. The OCR method's accuracy for the same work was measured 10 in count. Later for precision, sensitivity was also noted for both proposed and existing algorithms. With the aid of SPSS tool, performance comparison was carried out with independent samples t test and graph builder modules. The mean accuracy, mean precision and mean sensitivity were measured using proposed EOCR as 97.899%, 97.404% and 97.933% respectively and for OCR algorithm 94.059%, 94.925% and 95.387% respectively. The proposed EOCR method has gained more novelty to recognize images into title, subtitle and paragraph content than the OCR algorithm parentheses.

Keywords: Deep learning; machine learning; feature extraction; enhanced optical character recognition; visually impaired person assistance

DOI Number: 10.14704/nq.2022.20.12.NQ77047 **NeuroQuantology 2022; 20(12): 634-650**

Introduction

The exploration of ML a lgorithms plays an important part in the field of Artificial Intelligence (AI). Neural networks, a beautiful biologically- inspired programming paradigm which enables a computer to learn from experimental data including assistance for visually impaired persons [1]. Some kinds of hardware and software are being introduced for them as supportive assistance. These assistants can be physical device based assistance to some technical based assistance. One such work was

discussed about smart vision [2]. Here techniques such as Global position system (GPS), Ultrasonic Sensor, Radio Frequency Identification (RFID) etc. computer vision based object detection is an application for visually impaired persons to assistance like daily chores [3]. In this some DL and ML algorithms were applied on such assistance. The assistance such as text, sound and images and their features were introduced for the visually impaired. For this multiple abstraction and data representation were used.

According to the World Health Organization (WHO) statistics, there are nearly 285 million visually impaired people in the world, of which 140 million are elderly people over the age of 50, and 110 million of these visually impaired elderly suffer from multiple chronic diseases [4]. In recent times, several frameworks based on wearable devices and devoted to visually impaired people have surfaced. Visual impairment could be an individual's difficulties with traditional everyday activities such as driving, reading, socialization, and walking. It's important to advance devices with smart devices to assist visually-impaired people with these day-to-day activities [5].

Reading books is essential in acquiring information about various subjects. Visually impaired people have less access to information as many books are not transcribed in braille format. Thus, visually impaired people have to be dependent on someone even for a basic activity as simple as reading. Currently available methods for reading books are braille, scribes, Recorded book collections, Assistive Technology and eBook reader. There are only a few books that are transcribed into braille format and use of assistive technology is limited to eBooks. One such work was DL strategy which used braille for character recognition by visually impaired persons [6]. In this work, a light weight Convolutional Neural Network (CNN) was introduced to recognize character. This work has read accurate optical characters and hence performance was good. Dataset introduced here were Chinese double sided braille image and English braille. In general character recognition system is a technique, which has been recognized by difference DL algorithms including CNN, Recurrent Neural Networks (RNN), Regions with Convolutional Neural Networks (RCNN) etc. the applications of this image intro text documents were applied in assisted visually impaired person, character recognize for both printed and hand written character recognition systems etc.

Further this work has been outlined as follows. Section 2 has introduced the detailed description of related work of image into text. The implementation and its algorithm working mechanism have been introduced in Section 3. The results were discussed in Section 4 for performance measures such as accuracy, precision and sensitivity. The comparative discussion about existing methods and proposed method was elaborated in Section 5. Conclusion about proposed work's procedures and advantages and future work were discussed in Section 6 as end of this work.

2 Related Works on image Reading

In this section, it is essential to know different efficient approaches applied for converting imaged documents into textual one. A work was discussed with Adam optimization algorithm with help of DL. In this work to overcome the OCR as it has an inability to recognize handwritten characters supposed to be recognized. Alongside, the OpenCV package of Python was utilized to implement the Adam optimization algorithm. The dataset used for this algorithm was the Modified National Institute of Standards and Technology (MNIST) dataset with almost 60, 000 images of handwritten characters. For digital format of

image conversion from table formatted data image was carried out using OCR tool [8]. In this work extraction of rows and columns were considered for measuring performance of OCR. A mechanism for image into textual data was proposed with OCR for offline and online conversion especially for handwritten character recognition [9]. This work was developed as software, which has been utilized for more users over offline or online. Cursive content was perceived with implementing a framework. Error rate was measured and based on font size error rates were increased and decreased. Here character lines were determined based on individual letters and symbols. An encoded editable text processing technique was introduced with OCR alongside different Long Short-Term Memory (LSTM) libraries [10]. This work involved various datasets with different settings of OCR with data such as binarized and grey-scale image for testing and training alongside different effects. This work has introduced a framework called any OCR system. Some facts were analyzed to measure the performance of this framework. Character error rate and accuracy were measured to recognize for anyOCR method.

A work was introduced for neural word representation as entity recognition as information extraction [11]. This work was discussed only for visually rich banking documents and text intensive abide using OCR method. An interesting survey was introduced on OCR for image to text conversion [12]. This survey has listed different applications and methodologies on OCR image concepts with respect to companies and research for character recognition. It has detailed representation on pre-processing, image acquisition, segmentation of character, extraction of feature of character, classification alongside post-processing about OCR concept. A work has discussed character recognition with automation of text detection alongside local image features [13]. In this work technique called contour processing was introduced. For phrases and words it has been recognized from images using the OCR system. The content discovery has been aimed and hence it shows the image to word recognition.

A sequence neural networks based handwritten recognition system for continuous offline work completion was proposed [14]. The dataset such as RIMES and IAM were used with the LeNet-5 architecture of CNN. The identification pattern of handwritten character recognition was well established of considerable performance here. The technique was coined for encoding the visual characteristics as sequence to sequence architecture which was further decoded from images of handwritten characters. Error and accuracy analysis were introduced to measure the performances. Interestingly some researchers have introduced a handwritten recognition system with 3 Dimensional (3D) and it has been analyzed with some classifiers [15]. This work has differentiated fully from 2 Dimensional (2D) especially for cursive letters based image dataset. Bidirectional Long-Short Term Memory Neural Network (BLSTM-NN) was used as a method innovatively alongside the Recognizer Output Voting Error Reduction (ROVER) framework. The accuracy performance measure was measured for scripts such as Devanagari and Latin fonts. An Arabic OCR concept was introduced to recognize characters from holy dataset [16]. Here optimization technique was applied and the same was compared with mushaf al madinah benchmark. In this dataset, characters and diacritics were taken purely for optimization and accuracy. The proposed optimization based OCR was outperformed by the K-Nearest Neighbour (KNN) algorithm. For diacritics and characters, accuracy measures were analyzed. More time

complex process was the limitation of this work alongside the specialized dataset.

The real time images such as newspaper, magazine and captured were analysed and converted into smart text for visually impaired persons audio assistant and this work was introduced [17] recently. A smart work has been introduced for visually impaired persons converting image into speech with OCR for extraction of text further has been converted with guidance of eSpeak alongside audio amplifier [18]. Here Raspberry pi was utilised to interface image processing and keyboard. A natural scene detection and automated text detection solution was proposed for visually impaired persons [19]. For the process of this work, Contrast Limited Adaptive Histogram Equalization (CLAHE) was used to enhance the image from the dataset. Further Maximally Stable External Regions (MSER) used for feature detecting for text alongside filters for non-text identification. Finally OCR was used for text detection. The extended work of this solution used an audio synthesiser. A CNN based SqueezeNet was introduced with BatchNormalization as a special layer for visually impaired persons. This work has used Multi labeling of images into text [20]. This work has used a multi-labeling dataset with extended accuracy and time to process as performance measure. A module has been created to guide visually impaired people as Blindsys with deep CNN mechanisms. This work was introduced as object identification. Tab. 1 has listed various approaches for identification of text from images of different dataset and its limitations

.

Table 1: Comparative analysis of various approaches for text detection and their limitations

Name of work	Method	Utilized dataset	Limitation
Handwritten character recognition [7]	Adam optimization algorithm	MNIST dataset	More complex
Table data extraction [8]	OCR technique	Handwritten dataset	Less accurate prediction
OCR training	anyOCR	Archiv der DDR-Opposition	More complex and less accurate
Handwritten character recognition [14]	LeNet architecture	RIMES and IAM	More complex
3D handwritten character recognition [15]	BLSTM-NN	Devanagari and Latin image dataset	Less accuracy
Arabic character recognition [16]	OCR optimization technique	Characters and diacritics dataset	More complex structure
Smart reading glasses [18]	OCR with eSpeak	Real time image	Less efficient
Image to text and voice [19]	MSER with OCR	Standard natural scene	More complex
Presence of object in	SqueezeNet deep CNN	Multi-labelled	More complex

a scene [20]	dataset

Some research has been introduced for converting image to text followed by audio assistance to visually impaired persons, there was some lacunae being present. Since there was no proper solution to guide visually impaired people to get which part of the image of the paragraph has running sentences, subtitles, title etc. To help visually impaired people to overcome this difficulty, it is essential to propose a system in which the document is read electronically and will be converted to a multi-labeling feature. The ultimate aim of this work is to implement EOCR to convert images of any angle into text and guide visually impaired persons to get more appropriate reading assistance.

The OCR working mechanism was illustrated in Fig. 1. This consists of 6 blocks based on a set of activities. They were image input, acquisition of image, image digitization, image pre-processing, extraction / analysis of features and finally output of image.

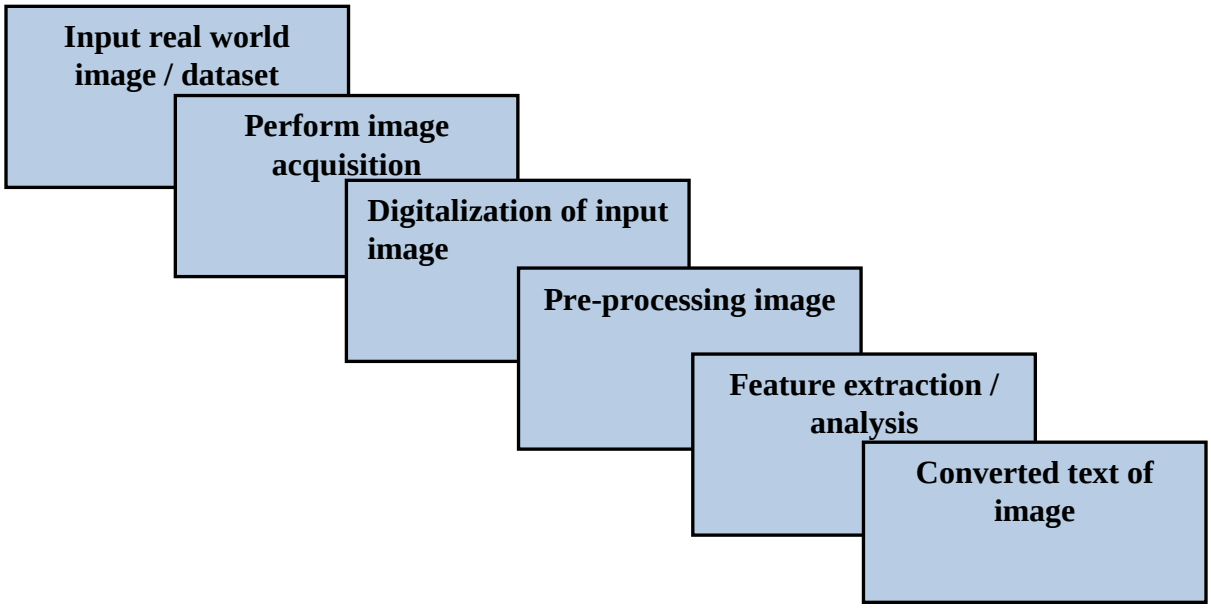

Figure 1: Steps of OCR technique for converting image to text

3 Material and Methods for Recognizing

This work was carried out to gain knowledge for research over conversion of image based text into visually impaired assist able textual with multi-section of text. Dataset was collected from the real world about 200 in count with different angles and rotations. 80% G power calculated with 0.05% error (alpha) with confidence interval 95%. Among 200, 80% was for training and 20% for testing.

3.1 Proposed EOCR algorithm's block diagram

Fig. 2 is showing the workflow of the EOCR method, which was used in this proposed work. Input image from dataset was trained and following test of dataset. This block diagram of EOCR has the following components. Here, the following explanation was being given for each component. EOCR has used the Tesseract module of python for this experimental setup as it was used under computer vision. Fig. 2 is showing the flow of EOCR and its working mechanism.

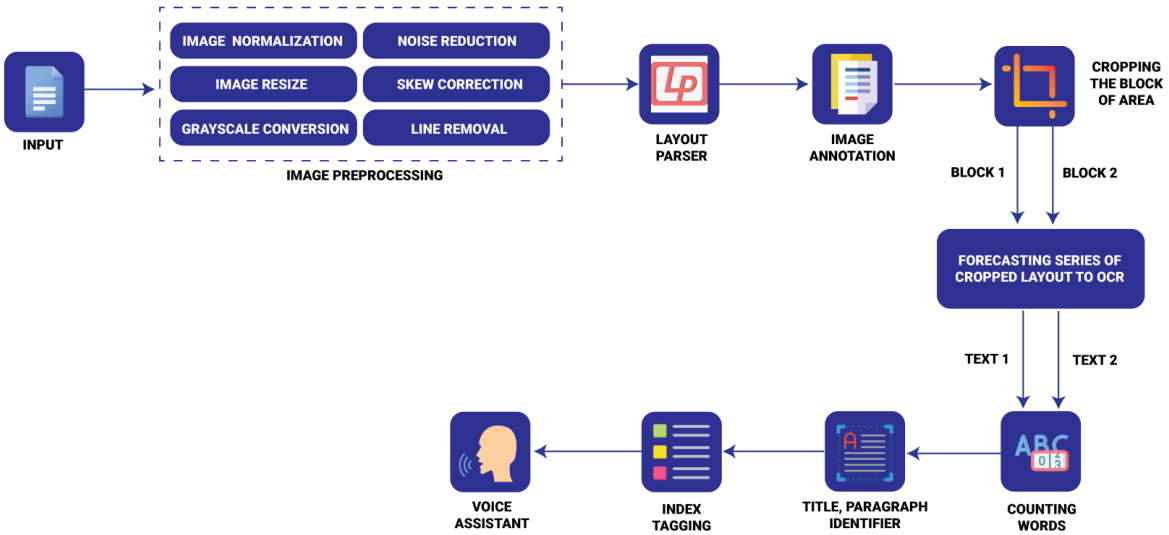

Figure 2: Block diagram for EOCR and its working mechanism

3.1.1 Input

This component was getting the input image either as jpg or png format.

3.1.2 Image pre-processing

This component has a couple pre-processing sub-components such as image normalization, noise reduction, image resize, skew correction, gray scale conversion, and finally line removal from the image. Usually these sets of processes in pre-processing of images would be carried out to make it a conversion ready procedure. Finally threshold value would be deciding defined size threshold and it is represented as Eq. (1).

$$C(i,j) = \frac{Imax - imin}{imax - imin + \varepsilon} \quad (1)$$

where C(i,j) is the threshold defined size in locality of image function, i_{max} is maximum pixel value and i_{min} is minimum pixel value.

3.1.3 Layout Parser

This component was used to detect layout of paper images further followed by parsing the output using coordinate system even for the complex image based layout and it was performed almost with DL concept. This is represented as Eq. (2).

$$db(i, l, fs, wbot, tbw) \quad (2)$$

where db is drawbox function, i is image, l is layout, fs is font size, wbot is with box on text, text box width.

3.1.4 Image Annotation

This component plays an inevitable role as follows. It initially prepares a dataset of image, followed by assigning object labels for the dataset of image. Further it draws the boundary box for objects supposed to detect. Finally it will classify the labels for each boundary box. It is represented as Eq. (3).

$$l = oagfta(r, al.GCVFT) \quad (3)$$

where l is layout, oagfta is OCR agent to gather full text annotation function, r is result, al is aggregate level and GCVFT is GCV Feature Type.

3.1.5 Cropping area of block

This component was used to crop the usable area of the image from the dataset. Also it detects the image into digitized text cleanly. Other functionality of this crop also used to reduce error rate and increase the accuracy of text conversion.

3.1.6 Forecasting series of cropped layout

This component can crop the image using layout function based on height and width of characters.

3.1.7 Word counting

This component was essential in image to text identification, since count of words will be deciding whether a box of words is title or sub-title or paragraph.

3.1.8 Identification of title/paragraph

Actually this component was introduced for enhanced one to convert OCR into EOCR. This intelligent component will decide, whether sequence of words were paragraph or title, based on count of words.

3.1.9 Index Tagging

This index will be used very sophisticatedly by visually impaired people, who can easily use the converted text as advanced assistance.

3.1.10 Voice Assistant

This will enable visually impaired people to gain knowledge of text of a particular page or anything.

3.2 The EOCR algorithm

The conversion of image into text has been carried out using the EOCR algorithm. This algorithm was working as represented in Algorithm 1. This algorithm has described the step by step process to convert an image into multi-feature based text format of each image page.

Algorithm 1: the EOCR algorithm

Step 1: Input the image of in jpeg, png format

Step 2: For every image in dataset, call Tesseract OCR engine

Step 3: Identify the page number, block number, paragraph number, confidence level, height, width

Step 4: Draw bound box and identify block number

Step 5: Crop the image as per height and width for each block number

Step 6: De-structure the original text image into blocks

Step 7: Store each blocks as jpeg format

Step 8: Load and pass again to OCR engine for text extraction

Step 9: Store the converted word and stored in the list data type

Step 10: Concatenate the list using join () function

Step 11: Identify paragraph or title using the if- else condition,

if (len (word_list [0])>10) then paragraph

else if (len (word_list[0])<10) then title

else empty

Step 12: Use confidence level for the classification

Step 13: Call read assistance using “gtts” library for converting the text into voice

3.3 Experiment setup and specification of system

The Jupiter version of the python tool was used with Tesseract OCR engine for conversion process of image into multi-featured text. For training and testing Algorithm 1 was called as it was coined as EOCR. 200 images were taken as dataset under ratio of 80% to 20% respectively for training and testing for this experiment. 10 counts of repeat were carried out to measure performance like accuracy, precision and sensitivity.

With conducted experiment alongside Algorithm 1 very accurate conversion of image to multi-featured text being obtained. This experimented output has produced multi-featured text such as identification of title, and ordinary paragraph. Further this feature would assist visually impaired persons to uniquely read which title has to be read or which paragraph has to be read, based on visually impaired persons interest and need. Fig. 3 is showing sample input and output. Actually the left hand side image was tested with Algorithm 1 and the respective output has been represented at the right side.

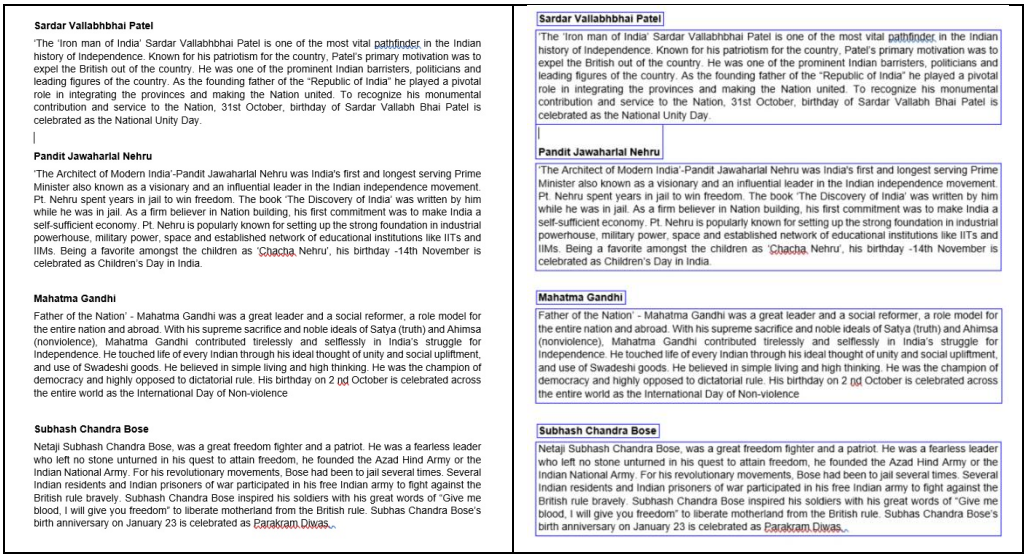

Figure 3: Identification of multi-featured text from sample input (left side-input, right side- output)

3.4Performance parameters for detection of image to text

In this experiment 3 performance measures were tried such as accuracy, precision and sensitivity. Usually confusion matrix would be having raw informative parameters such as True Positive (TP) which is nothing but taking image as input and respective true text identified as output, True Negative (TN) which is nothing but taking image as input and respective output getting wrong text, False Positive (FP) which is nothing but taking image as wrong image and getting true text and finally False Negative (FN) which is nothing but taking wrong input and getting wrong output. This information was represented in Tab. 2.

Table 2: Image to text detection confusion matrix

Identification Approach	Has true text	Does not have true text
Identified as true text	TP	FP

Not Identified as true text	FN	TN

Accuracy for right text identification is the right text captured from image and it rests on specific value closeness to expect. It is the ratio between the entirety of TP and TN with the quantity of TP, TN, FP and FN together. The computing accuracy is devised as Eq. (4).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

Precision of text detection is the immediacy extent of finding and it is the percentage between TP with amount of TP and FP. Eq. (5) is invented for determining precision.

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

Finally sensitivity performance measure was calculated, it is nonentity but, the shortest intention of TP of text detection to the ancillary intention of text detection's entirety of TP and FN. The sensitivity is characterized as Eq. (6).

$$Sensitivity = \frac{TP}{TP+FN} \quad (6)$$

4 Results about efficient recognition

The multi-featured text conversion from image dataset was carried out with EOCR and OCR algorithm. For this 80% dataset was used for training and 20% dataset was utilized for testing. Further SPSS tool utilized to analyze thrown results of experiment.

4.1 Statistical analysis

The experimented performance was measured as accuracy, precision and sensitivity for image to text for guiding visually impaired persons on textual reading. This experiment gained multi-feature based text. Same wise performance measures were noted on 10 counts. These performance measures were applied on the SPSS tool to initiate analysis to claim standard deviation and standard error rates. These experimented results were used modules of SPSS like independent t sample test and graphics builder. These modules were strongly utilized to compare accuracy, precision and sensitivity of both OCR and EOCR algorithms.

4.2 Performance measures of ECOR algorithm

Tab. 3 is listing the performance parameters values such as accuracy, precision and sensitivity 10 in count. These values have been measured with help of conducted experiment between EOCR and OCR for multi-feature text conversion from image dataset including abnormal images.

Table 3: Performance value for accuracy, precision and sensitivity between EOCR and OCR algorithms

Algorithm	Accuracy	Precision	Sensitivity	Algorithm	Accuracy	Precision	Sensitivity
	97.76	98.43	98.13		90.65	91.69	90.98
	97.26	97.17	98.25		92.36	92.57	93.67
	98.22	95.14	97.41		94.45	93.89	94.76
EOCR	96.83	97.87	98.62	OCR	93.13	94.88	95.72
	98.74	98.28	98.25		93.22	95.79	96.97
	98.14	97.23	96.45		94.59	96.57	97.45
	97.37	96.82	97.87		94.20	95.56	95.53

	98.62	98.75	98.71	96.28	97.82	95.89
	98.77	97.52	98.26	97.23	94.89	96.48
	97.28	96.83	97.38	94.48	95.59	96.42

4.3 Comparison of accuracy EOCR with OCR methods

With guidance of the SPSS tool, the mean accuracy was measured along with standard error rate and standard deviation for EOCR and OCR algorithms. It was inferred that EOCR mean accuracy as 97.899% and which was considerably better than the mean accuracy of OCR, which was inferred as 94.059% as presented in Tab. 4. It was also inferred that there was considerable high difference between standard error mean and standard deviation between EOCR and OCR algorithm with respect to mean accuracy comparison.

Table 4: Mean accuracy comparison between EOCR and OCR algorithms with standard deviation

Group Statistics					
	Algorithm	N	Mean	Std. Deviation	Std. Error Mean
Accuracy	EOCR	10	97.899	0.698	0.221
	OCR	10	94.059	1.877	0.593

Tab. 5 was listing the independent sample test parameters' values for mean accuracy comparison between EOCR and OCR algorithm, which was actually generated with guidance of Statistical analysis of SPSS under independent t sample test module. Here the notable inference has been noted that there was significant difference between EOCR and OCR gained as 0.061% which almost equal to 0.05 (error alpha value) and hence EOCR algorithm is better than OCR algorithm for mean accuracy. Other parameters like F, t, df, mean difference, standard error difference also have better comparison between EOCR and OCR. Thus mean accuracy of EOCR has attained novel performance for identification of multi-featured text from image dataset. Indirectly this multi-feature would assist visually impaired people to read with voice assistance.

Table 5: Independent sample t test for accuracy comparison between EOCR and OCR algorithm

Independent Samples Test										
		Lev. Test for Eq. of Var.		t-test for Eq. of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Diff.	Std. Err. Diff.	95% Conf. Int. of the Diff.	
									Lower	Upper
Accuracy	Eq. var. assumed	3.99	0.061	6.06	18	0.00	3.84	0.63	2.56	5.17
	Eq. var. not assumed			6.06	11.44	0.00	3.84	0.63	2.45	5.22

The simple bar mean graph has been generated with guidance of the graph builder module. Here the x – axis was assigned with algorithm types such as EOCR and OCR and the y-axis was assigned with mean accuracy values of EOCR and OCR. The standard deviation was set as ±1 and confidence interval as 95%. Graph was generated under this state and the generated graph was shown in Fig. 4. The inference was that EOCR algorithm has gained more accuracy to detect multi-feature text efficiently than OCR

algorithm as mean accuracy was considerably lower for OCR. Thus novelty has been introduced with help of EOCR as it has retained remarkable performance for mean accuracy.

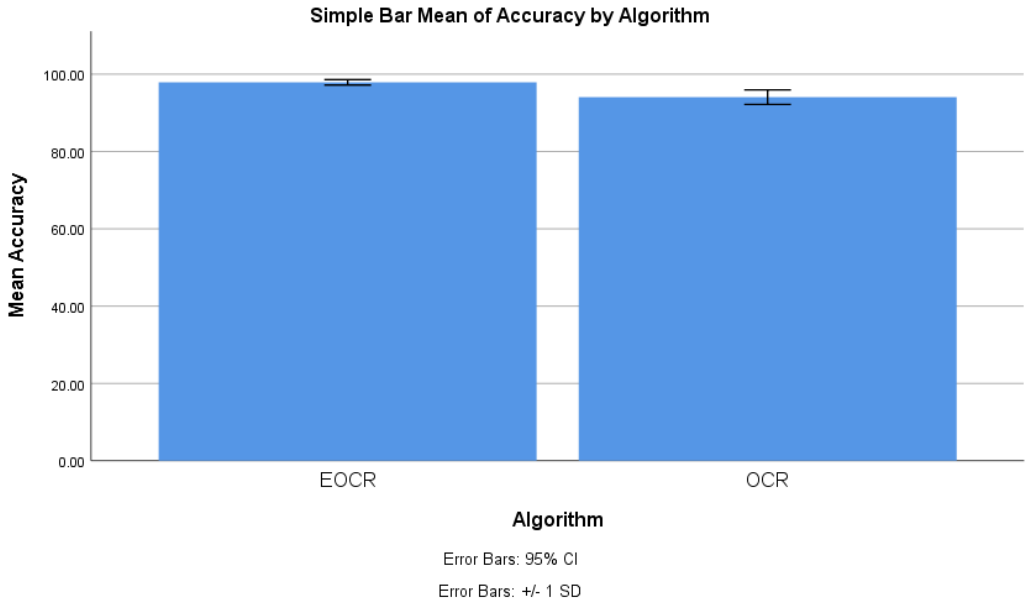

Figure 4: Mean accuracy comparison between EOCR and OCR for multi-feature text conversion

4.4 Comparison of precision EOCR with OCR methods

With guidance of the SPSS tool, the mean precision was measured along with standard error rate and standard deviation for EOCR and OCR algorithms. It was inferred that EOCR mean precision as 97.404% and which was considerably better than the mean precision of OCR, which was inferred as 94.925% as presented in Tab. 6. It was also inferred that there was considerable high difference between standard error mean and standard deviation between EOCR and OCR algorithm with respect to mean precision comparison.

Table 6: Mean precision comparison between EOCR and OCR algorithms with standard deviation

Group Statistics					
	Algorithm	N	Mean	Std. Deviation	Std. Error Mean
Precision	EOCR	10	97.404	1.041	0.329
	OCR	10	94.925	1.819	0.575

Tab. 7 was listing the independent sample test parameters' values for mean precision comparison between EOCR and OCR algorithm, which was actually generated with guidance of Statistical analysis of SPSS under independent t sample test module. Here the notable inference has been noted that there was significant difference between EOCR and OCR gained as 0.185% which slightly more to 0.05 (error alpha value) and hence EOCR algorithm is better than OCR algorithm for mean precision. Other parameters like F, t, df, mean difference, standard error difference also have better comparison between EOCR and OCR. Thus mean precision of EOCR has attained novel performance for identification of multi-featured text from image dataset.

Table 7: Independent sample t test for precision comparison between EOCR and OCR algorithm

Independent Samples Test

		Lev. Test for Eq. of Var.		t-test for Eq. of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Diff.	Std. Err. Diff.	95% Conf. Int. of the Diff.	
									Lower	Upper
Precision	Eq. var. assumed	1.90	0.185	3.741	18	0.001	2.479	0.663	1.087	3.871
	Eq. var. not assumed			3.741	14.32	0.002	2.479	0.663	1.061	3.897

The simple bar mean graph has been generated with guidance of the graph builder module. Here the x – axis was assigned with algorithm types such as EOCR and OCR and the y-axis was assigned with mean precision values of EOCR and OCR. The standard deviation was set as ±1 and confidence interval as 95%. Graph was generated under this state and the generated graph was shown in Fig. 5. The inference was that EOCR algorithm has gained more precision to detect multi-feature text efficiently than OCR algorithm as mean precision was considerably lower for OCR. Thus novelty has been introduced with help of EOCR as it has retained remarkable performance for mean precision.

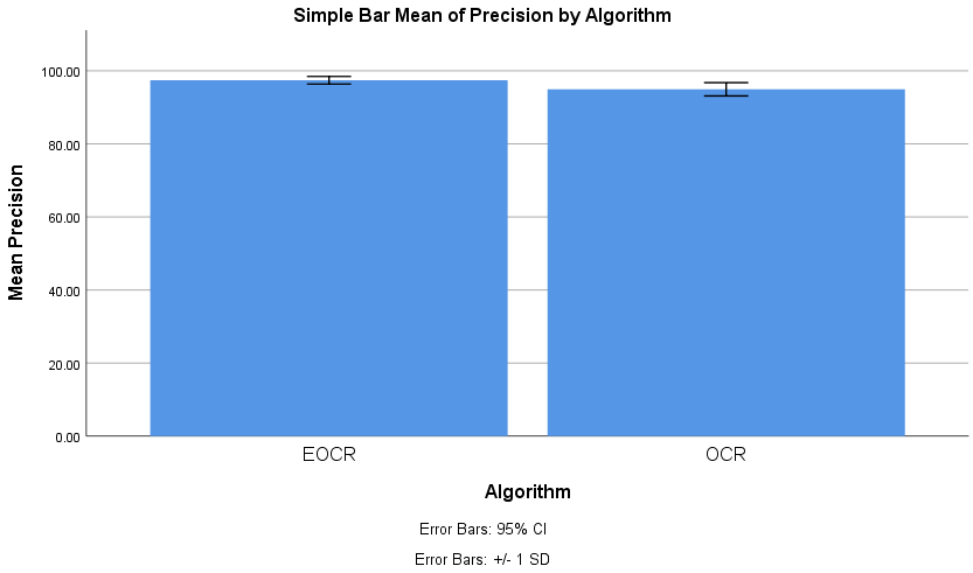

Figure 5: Mean precision comparison between EOCR and OCR for multi-feature text conversion

4.5 Comparison of sensitivity EOCR with OCR methods

With guidance of the SPSS tool, the mean sensitivity was measured along with standard error rate and standard deviation for EOCR and OCR algorithms. It was inferred that EOCR mean sensitivity as 97.933% and which was considerably better than the mean sensitivity of OCR, which was inferred as 95.387% as presented in Tab. 8. It was also inferred that there was considerable high difference between standard error mean and standard deviation between EOCR and OCR algorithm with respect to mean sensitivity comparison.

Table 8: Mean sensitivity comparison between EOCR and OCR algorithms with standard deviation

Group Statistics

	Algorithm	N	Mean	Std. Deviation	Std. Error Mean
Sensitivity	EOCR	10	97.933	0.684	0.216
	OCR	10	95.387	1.891	0.598

Tab. 9 was listing the independent sample test parameters' values for mean sensitivity comparison between EOCR and OCR algorithm, which was actually generated with guidance of statistical analysis of SPSS under independent t sample test module. Here the notable inference has been noted that there was significant difference between EOCR and OCR gained as 0.061% which slightly more to 0.05 (error alpha value) and hence EOCR algorithm is better than OCR algorithm for mean sensitivity. Other parameters like F, t, df, mean difference, standard error difference also have better comparison between EOCR and OCR. Thus mean sensitivity of EOCR has attained novel performance for identification of multi-featured text from image dataset.

Table 9: Independent sample t test for sensitivity comparison between EOCR and OCR algorithm

Independent Samples Test										
		Lev. Test for Eq. of Var.		t-test for Eq. of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Diff.	Std. Err. Diff.	95% Conf. Int. of the Diff.	
									Lower	Upper
Accuracy	Eq. var. assumed	3.98	0.061	4.005	18	0.001	2.546	0.636	1.210	3.882
	Eq. var. not assumed			4.005	11.31	0.002	2.546	0.636	1.151	3.941

The simple bar mean graph has been generated with guidance of the graph builder module. Here the x – axis was assigned with algorithm types such as EOCR and OCR and the y-axis was assigned with mean sensitivity values of EOCR and OCR. The standard deviation was set as ±1 and confidence interval as 95%. Graph was generated under this state and the generated graph was shown in Fig. 6. The inference was that EOCR algorithm has gained more sensitivity to detect multi-feature text efficiently than OCR algorithm as mean sensitivity was considerably lower for OCR. Thus novelty has been introduced with help of EOCR as it has retained remarkable performance for mean sensitivity.

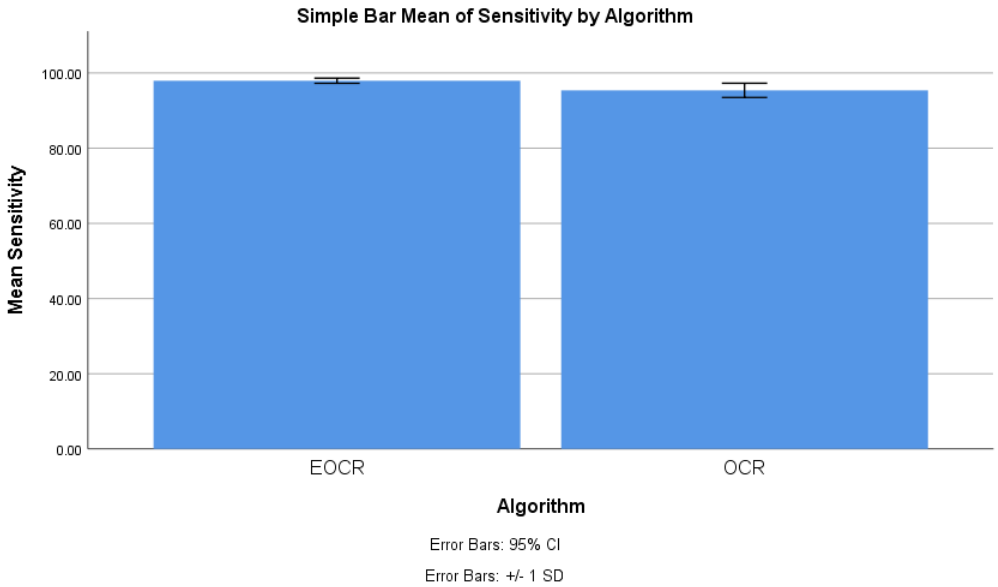

Figure 6: Mean sensitivity comparison between EOCR and OCR for multi-feature text conversion

4.6 Overall performance of results over EOCR with OCR methods

Tab. 10 has listed the cumulative performance measures values of accuracy, precision and sensitivity between EOCR and OCR algorithm for multi-feature text detection from real time image dataset. Cumulatively EOCR has measured its performance more than the OCR algorithm.

Table 10: Comparative performance analysis between EOCR and OCR algorithms

Group Statistics					
	Algorithm	N	Mean	Std. Deviation	Std. Error Mean
Accuracy	EOCR	10	97.899	0.698	0.221
	OCR	10	94.059	1.877	0.593
Precision	EOCR	10	97.404	1.041	0.329
	OCR	10	94.925	1.819	0.575
Sensitivity	EOCR	10	97.933	0.684	0.216
	OCR	10	95.387	1.891	0.598

The simple bar mean graph has been generated with guidance of the graph builder module. Here the x – axis was assigned with algorithm types such as EOCR and OCR and the y-axis was assigned with cumulative values of EOCR and OCR performance parameters. The standard deviation was set as ±1 and confidence interval as 95%. Graph was generated under this state and the generated graph was shown in Fig. 7. The inference was that EOCR algorithm has gained more cumulative performance to detect multi-feature text efficiently than OCR algorithm as cumulative performance was considerably lower for OCR. Thus novelty has been introduced with help of EOCR as it has retained remarkable cumulative performance.

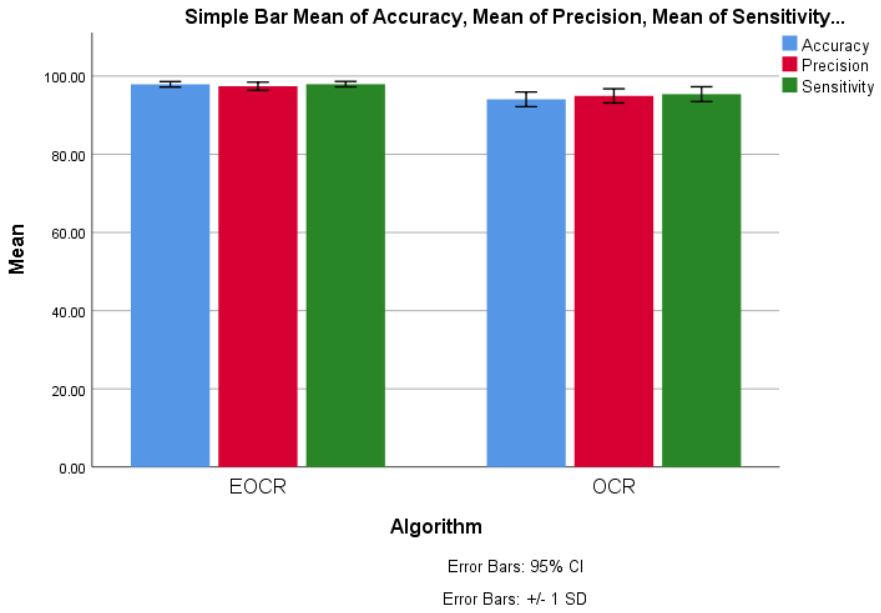

Figure 7: Cumulative performance comparison between EOCR and OCR for multi-feature text

5 Discussion on Inference of Experiment

Here the inference was noted from the graphics builder module and independent sample test for accuracy performance. It was inferred that proposed EOCR algorithm has gained maximum accuracy as 98.77% with real time dataset. To compare similar and dissimilar findings, here the comparison has been deployed with existing ML and DL based algorithms for detecting text from image with OCR principle. Tab. 11 has listed the accuracy comparison between various existing work's accuracies with proposed EOCR algorithm accuracy. It was noted that CNN+SVM has introduced accuracy results as 94.40%, while CNN+KNN has gained accuracy as 96.98% [7]. One version of OCR called anyOCR has produced accuracy to detect text from images as 84.6% [10], which was comparatively less than the proposed method. A work used BLSTM-NN as DL method to detect text and has introduced accuracy as 84.6% [15]. The method called OCR optimization technique has produced accuracy for text detection as 92.301% [16], which was almost lagging with 6.4% lesser than proposed EOCR algorithm accuracy. This comparison has concluded that similar finding and dissimilar finding were almost lagging with their accuracy while comparing for text detecting from image dataset. Hence this discussion part brings the provision that the proposed algorithm would be a better suggestion in detecting text accurately, precisely and sensitivity than existing algorithms.

Table 11: Accuracy comparison of EOCR with existing techniques for text detection from image

Method	Dataset	Accuracy
CNN+SVM, CNN + KNN [7]	MNIST	94.40%, 96.98%
anyOCR [10]	Archiv der DDR-Opposition	84.6%
BLSTM-NN [15]	Devanagari and Latin image	72.25%
OCR optimization technique [16]	Characters and diacritics	92.301%
Proposed EOCR algorithm	Real time dataset	98.77%

6 Conclusion

Aiding visually impaired people to read text based content was challenging before technology was

constructed. Nowadays different methods have been introduced to guide visually impaired people to read text, either directly by themselves or with another person or technological assistant. Using ML and DL, some work has introduced solutions to aid visually impaired people. But lagging constraints were present. To overcome such omission, in this research work, the EOCR algorithm has been introduced to detect multi-featured text to aid visually impaired persons to read text as per convenience. The experiment was conducted to measure the performance parameters such as accuracy, precision and sensitivity to detect multi-featured text to assist visually impaired persons. These performance measures have been analyzed with guidance of SPSS statistical tools. The inferences claimed that the EOCR algorithm has performed well with performance parameters significantly. The mean accuracy, mean precision and mean sensitivity were analyzed and their values were introduced as 97.899%, 97.404% and 97.933% respectively for EOCR algorithm, which were significantly better than OCR algorithm. To work out more aid for visually impaired persons, in future it was decided to extend this work for guiding identification of multiple objects from video and the same has to be guided to recognize [21] [22] with Enhanced DL and ML algorithms [23] [24].

References

[1] T. Priya, K.S. Sravya and S. Umamaheswari, “Machine-learning-based device for visually impaired person,” in Artificial Intelligence and Evolutionary Computations in Engineering Systems Advances in Intelligent Systems and Computing, vol. 1056. Singapore: Springer, pp. 79-88, 2020. [Online]. Available: https://doi.org/10.1007/978-981-15-0199-9_7

[2] K. G. Shanthi “Smart vision using machine learning for blind,” International Journal of Advanced Science and Technology, vol. 29, no. 5, pp. 12458-12463, 2020.

[3] P. Chandankhede and A. Kumar, "Deep learning technique for serving visually impaired person," in Proc. ICETET-SIP, Nagpur, India, pp. 1-6, 2019.

[4] J. L. Wolff, B. Starfield and G. Anderson “Prevalence, expenditures, and complications of multiple chronic conditions in the elderly,” Archives Internal Medicine, vol. 162, no. 20, pp. 2269–2276, 2002.

[5] Y. S. Su, C. H. Chou, Y. L. Chu and Z. Y. Yang, "A finger-worn device for exploring chinese printed text with using CNN algorithm on a micro IoT processor," IEEE Access, vol. 7, pp. 116529-116541, 2019.

[6] K. Smelyakov, A. Chupryna, D. Yeremenko, A. Sakhon and V. Polezhai, "Braille character recognition based on neural networks," in Prod. DSMP, Lviv, Ukraine, pp. 509-513, 2018.

[7] H. Kohli, J. Agarwal and M. Kumar, "An improved method for text detection using adam optimization algorithm," Global Transitions Proceedings, vol. 3, pp. 230-234, 2022.

[8] M. Y. Akpinar, E. Emeklıgıl and S. Arslan, "Extracting table data from images using optical character recognition text," in Proc. SIU, Izmir, Turkey, pp. 1-4, 2018.

[9] R. Satheeskumar, “Optical and Intelligent Character Recognition System,” International Journal of Creative Research Thoughts, vol. 10, no. 3, pp. 957-968, 2022.

[10] S. S. Bukhari, S. Francis, C. N. N. Kamath and A. Dengel, "An Investigative Analysis of Different LSTM Libraries for Supervised and Unsupervised Architectures of OCR Training," in Proc. ICFHR, Niagara Falls, NY, USA, pp. 447-452, 2018.

eISSN 1303-515

[11] B. Oral, E. Emekligil, S. Arslan and G. Eryigit, "Information extraction from text intensive and visually rich banking documents," Information Processing & Management, vol.57, no.6, pp. 102361(1-24), 2020.

[12] N. Islam, Z. Islam and N. Noor, "A survey on optical character recognition system," Journal of Information & Communication Technology, vol. 10, no. 2, pp. 1-4, 2016.

[13] B. Remigiusz, P. Pavol and W. Rafal, “Automated text detection and character recognition in natural scenes based on local image features and contour processing techniques,” Intelligent Human Systems Integration, vol. 722, pp. 42-48, 2018.

[14] J. Sueiras, V. Ruiz, A. Sanchez and J. F. Velez, “Offline continuous handwriting recognition using sequence to sequence neural networks,” Neurocomputing, vol. 289, pp. 119-128, 2018.

[15] P. Kumar, R. Saini, P. P. Roy and U. Pal, “A lexicon-free approach for 3D handwriting recognition using classifier combination,” Pattern Recognition Letters, vol. 103, pp. 1-7, 2018.

[16] F. Alotaibi, M. T. Abdullah, R. B. H. Abdullah, R. W. B. O. K. Rahmat, I. A. T. Hashem et al., "Optical character recognition for quranic image similarity matching," IEEE Access, vol. 6, pp. 554-562, 2018.

[17] S. Manoharan, "A smart image processing algorithm for text recognition, information extraction and vocalization for the visually challenged,” Journal of Innovative Image Processing, vol. 1, no. 1, pp. 31-38, 2019.

[18] G. H. Nayana, N. Sowmya, Y. Sravani and B. James, "Smart reading glasses: conversion of image text into speech," International Journal of Creative Research Thoughts, vol. 6, no. 2, pp. 1515-1520, 2018.

[19] S. M. Qaisar, R. Khan and N. Hammad, "Scene to text conversion and pronunciation for visually impaired people," in Proc. ASET, Dubai, United Arab Emirates, pp. 1-4, 2019.

[20] H. Alhichri, Y. Bazi, N. Alajlan and B. B. Jdira, "Helping the visually impaired see via image multi-labeling based on SqueezeNet CNN," Application Science, vol. 9, no. 21, pp. 4656, 2019.

[21] M. Adimoolam, S. Mohan, A. John and G. Srivastava, ”A Novel technique to detect and track multiple objects in dynamic video surveillance systems,” International Journal of Interactive Multimedia and Artificial Intelligence, vol. 7, no. 4, pp. 112-120, 2022.

[22] N. M. Balamurugan, T. K. S. Rathishbabu, M. Adimoolam and A. John, ”A novel efficient algorithm for duplicate video comparison in surveillance video storage systems,” Journal of Ambient Intelligence and Humanized Computing, vol. 1, pp. 1-16, 2021.

[23] N. M. Balamurugan, M. Adimoolam, M. H. Alsharif, and P. Uthansakul, “A novel method for improved network traffic prediction using enhanced deep reinforcement learning algorithm,” Sensors, vol. 22, no. 13, pp. 5006 (1-17), 2022.

[24] N. M. Balamurugan, K. Maithili, T. K. S. R. Babu and M, Adimoolam, “Stage wise categorization and prediction of diabetic retinopathy using ensemble learning and 2D-CNN,” Intelligent Automation & Soft Computing, pp. 1-16 2022.