AHMED ANWAR <ahmed.anwar@g.bracu.ac.bd>

6th International Conference on Electrical Information and Communication Technology 2023 : Submission (28) has been created.

1 message

Microsoft CMT <email@msr-cmt.org>   Sun, May 14, 2023 at 9:34 PM
Reply-To: Microsoft CMT - Do Not Reply <noreply@msr-cmt.org>
To: ahmed.anwar@g.bracu.ac.bd

```
Hello,

The following submission has been created.

Track Name: Artificial Intelligence, Computer Networks, and Security

Paper ID: 28

Paper Title: Utilizing Optical Character Recognition to Summarize Bengali Text

Abstract:
Optical Character Recognition (OCR) has proven to be a valuable tool for extracting textual information
from images, converting physical documents into digital formats that can be easily edited, searched, and
shared. However, OCR technology is not without its limitations, particularly when it comes to recognizing
and processing non-Latin scripts such as Bengali. Over 230 million people speak Bengali, making it one of
the most widely used languages in the world. Summarizing Bengali material is one area where OCR could be
especially helpful, saving time and effort for scholars and journalists who are working with a big corpus
of text. In recent years, there have been several attempts to develop OCR-based text summarization
techniques for Bengali. One challenge it struggles with is the complex nature of the language, its rich
vocabulary, its highly inflected grammar and the inclusion of a number of ligatures in Bengali Script and
diacritics that can significantly complicate the OCR process. Overall, the development of OCR-based text
summarization could save time and effort, while also providing a more objective and systematic approach to
analyzing large volumes of text. It is expected that in the future, Bengali text summarizing methods based
on OCR will become more advanced and precise as technology advances.

Created on: Sun, 14 May 2023 15:34:26 GMT

Last Modified: Sun, 14 May 2023 15:34:26 GMT

Authors:
        - abhijit.saha@g.bracu.ac.bd (Primary)
        - faiza.bushra@g.bracu.ac.bd
        - ahmed.anwar@g.bracu.ac.bd
        - rubaba.rashid@g.bracu.ac.bd
        - ehsanur.rahman.rhythm@g.bracu.ac.bd
        - md.sabbir.hossain1@g.bracu.ac.bd
        - annajiat@gmail.com

Secondary Subject Areas: Not Entered
Submission Files:     Utilizing Optical Character Recognition to Summarize Bengali Text.pdf (283 Kb, Sun,
14 May 2023 15:33:52 GMT)

Submission Questions Response: Not Entered

Thanks,
CMT team.
```

```
Download the CMT app to access submissions and reviews on the move and receive notifications:
https://apps.apple.com/us/app/conference-management-toolkit/id1532488001
```

https://play.google.com/store/apps/details?id=com.microsoft.research.cmt

To stop receiving conference emails, you can check the 'Do not send me conference email' box from your
User Profile.

Microsoft respects your privacy. To learn more, please read our Privacy Statement.

Microsoft Corporation
One Microsoft Way
Redmond, WA 98052