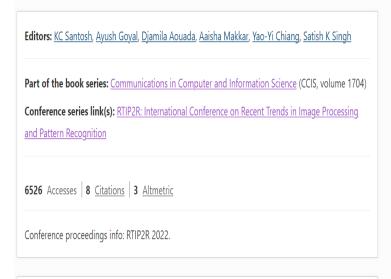


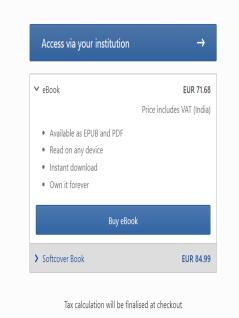
Conference proceedings © 2023

Recent Trends in Image Processing and Pattern Recognition

5th International Conference, RTIP2R 2022, Kingsville, TX, USA, December 1-2, 2022, Revised Selected Papers

$\underline{\mathsf{Home}}$ > Conference proceedings







The Classification of Native and Invasive Species in North America: A Transfer Learning and Random Forest Pipeline

Sayani Sarkar, Somenath Chakroborty Pages 297-307

Internet of Things and Security

Front Matter Pages 309-309 <u>PDF</u> **±**

Towards a Digital Twin Integrated DLT and IoT-Based Automated Healthcare

Prodipta Promit Mukherjee, Maharin Afroj, Sohaima Hossain, Milon Biswas

<u>Enabling Edge Devices Using Federated Learning and Big Data for Proactive Decisions</u>

Abishi Chowdhury, A. Swaminathan, Rajan R. Ashoka, Amrit Pal

<u>IoT and Blockchain Oriented Gender Determination of Bangladeshi Populations</u>

Md.Akkas Ali, Rajesh Kumar Dhanaraj Pages 337-343

Federated Learning Based Secured Computational Offloading in Cyber-Physical IoST Systems

A Hybrid Campus Security System Combined of Face, Number-Plate, and Voice Recognition

Abu Sayeed, Azmain Yakin Srizon, Md. Mehedi Hasan, Jungpil Shin, Md. Al Mehedi Hasan, M. Rasel Mahmud Pages 356-368

Signal Processing and Machine Learning

Front Matter Pages 369-369 <u>PDF</u> **±**

Single-Trial Detection of Event-Related Potentials with Artificial Examples Based on Coloring Transformation

Hubert Cecotti, Steve Jaimes Pages 371-382

<u>Identifying the Relationship Between Hypothesis and Premise</u>

Srishti Jhunthra, Harshit Garg, Vedika Gupta

Data Poisoning Attack by Label Flipping on SplitFed Learning

Saurabh Gajbhiye, Priyanka Singh, Shaifu Gupta

A Deep Learning-Powered Voice-Enabled Math Tutor for Kids

Arnab Banerjee, Srijoy Paul, Tisu Priya, Anamika Rohit, Nibaran Das

Back to top ↑

About this book

This book constitutes the refereed proceedings of the 5th International Conference on Recent Trends in Image Processing and Pattern Recognition, RTIP2R 2022, held in Kingsville, TX, USA, in collaboration with the Applied AI Research Laboratory of the University of South Dakota, during December 01-02, 2022.

The 31 full papers included in this book were carefully reviewed and selected from 69 submissions. They were organized in topical sections as follows: healthcare: medical imaging and informatics; computer vision and pattern recognition; internet of things and security; and signal processing and machine learning.

Back to top 1

Keywords

artificial intelligence classification methods computer security

computer vision correlation analysis cryptography databases

Editors and Affiliations

University of South Dakota, Vermillion, USA

KC Santosh

Texas A&M University, College Station, USA

Ayush Goyal

University of Luxembourg, Luxembourg

Djamila Aouada

University of Derby, Derby, UK

Aaisha Makkar

University of Minnesota, Minneapolis, USA

Yao-Yi Chiang

IIIT Allahabad, Allahabad, India

Satish K Singh

Back to top ↑

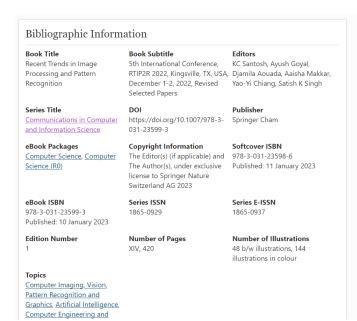
Bibliographic Information

Book Title

Book Subtitle

Editors

Recent Trends in Image Processing and Pattern 5th International Conference, KC Santosh, Ayush Goyal, RTIP2R 2022, Kingsville, TX, USA, Djamila Aouada, Aaisha Makkar,



A Deep Learning-Powered Voice-Enabled Math Tutor for Kids

<u>Arnab Banerjee</u> [™], <u>Srijoy Paul</u>, <u>Tisu Priya</u>, <u>Anamika Rohit</u> & <u>Nibaran Das</u>

Conference paper | First Online: 11 January 2023

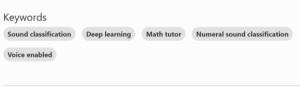
196 Accesses

Part of the Communications in Computer and Information Science book series (CCIS, volume 1704)

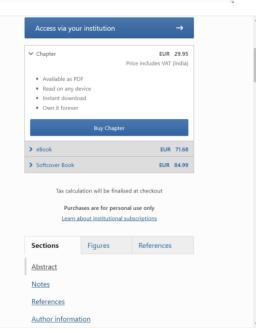
Abstract

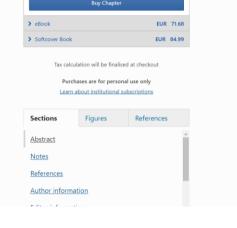
In this study, a voice enabled math tutor system is proposed that enables children to practice math problems on their own. For this, we have developed numerical sound dataset targeting the application. In the application, when the system is turned on, a math problem is generated, and the child will respond verbally to it. The system will categorize the audio data (user-provided answer) and produce a text number, which will then be further analysed and generate output as either a correct or erroneous answer. Any toy can be equipped with the proposed system, allowing a kid to practice problems while engaging with the system. A dataset named JUDVLP-BCRP: numeralSound.v1 is prepared, with 2315 audio data of numerals in the range of 0 to 9. In a typical setting, the audio data were collected from people in the age range of 10 to 60 from West Bengal, Jharkhand, Delhi, Assam, Bihar, and Orissa. After preprocessing the audios, Mel spectrograms were produced which acts as input by the deep neural network algorithms. The audio data has been classified using a number of well-known

deep learning algorithms, including DenseNet-121, VGG-16, modified DenseNet121 (DenseNet-41), and modified VGG-16 (VGG-12). Using DenseNet-121, VGG-16, DenseNet-41, VGG-12, 94.60%, 98.70%, 98.27%, and 98.48% accuracy was obtained. The networks were run for 100 epochs using a learning rate of 0.0001, and categorical cross-entropy loss function. The VGG-16 produced the highest precision of 98.9%, and the VGG-12 produced the second-best precision of 98.6%. The outcomes are positive and influence a workable system design.



Supported by Dr. B. C. Roy Polytechnic and Jadavpur University.





Author information Authors and Affiliations Dr. B. C. Roy Polytechnic, Durgapur, 713206, West Bengal, India Arnab Banerjee, Srijoy Paul, Tisu Priya & Anamika Rohit Jadavpur University, Kolkata, 700032, West Bengal, India Arnab Banerjee & Nibaran Das Corresponding author Correspondence to Arnab Banerjee . Editor information **Editors and Affiliations** University of South Dakota, Vermillion, SD, USA KC Santosh Texas A&M University, College Station, TX, USA Ayush Goyal University of Luxembourg, Luxembourg Djamila Aouada

