

Smart sensing-based approaches for Real-time River Water Pollution Monitoring

वास्तविक समय नदी जल प्रदूषण निगरानी के लिए स्मार्ट सेंसिंग-आधारित दृष्टिकोण



Thesis submitted in partial fulfillment
for the Award of Degree

Doctor of Philosophy

by

Swati Sandeep Chopade

स्वाति संदीप चोपड़े

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY
(BANARAS HINDU UNIVERSITY)
VARANASI - 221005

Roll No. 18071015

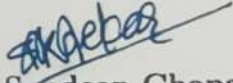
Year 2024

DECLARATION BY THE CANDIDATE

I, **Swati Sandeep Chopade**, certify that the work embodied in this Ph.D. thesis is my own bonafide work carried out by me under the supervision of **Dr. Hari Prabhat Gupta** from **July 2018** to **October 2023** at **Department of Computer Science and Engineering**, Indian Institute of Technology (BHU) Varanasi. The matter embodied in this thesis has not been submitted for the award of any other degree/diploma. I declare that I have faithfully acknowledged and given credits to the research workers wherever their works have been cited in my work in this thesis. I further declare that I have not willfully copied any other's work, paragraphs, text, data, results, *etc.* reported in journals, books, magazines, reports, dissertations, theses, *etc.*, or available at websites and have not included them in this thesis and have not cited as my own work.


Date:

Place: Varanasi


(Swati Sandeep Chopade)

CERTIFICATE BY THE SUPERVISOR

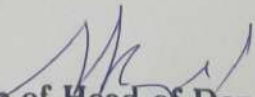
This is to certify that the above statement made by the candidate is correct to the best of my knowledge.


(Dr. Hari Prabhat Gupta)

Associate Professor,

Dept. of Computer Science and Engineering,

Indian Institute of Technology (BHU) Varanasi


Signature of Head of Department
आचार्य व विभागाध्यक्ष

Professor & Head

समूहक विज्ञान एवं अभियांत्रिकी विभाग

Department of Computer Sc. & Engg

भारतीय प्रौद्योगिकी संस्थान

Indian Institute of Technology

(बनारस हिन्दू यूनिवर्सिटी)

(Banaras Hindu University)

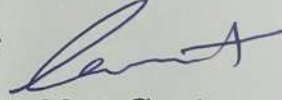
बनारस- 221005 / Varanasi-221005

CERTIFICATE

It is certified that the work contained in the thesis titled "*Smart sensing-based approaches for Real-time River Water Pollution Monitoring*" by *Swati Sandeep Chopade* has been carried out under my supervision and that this work has not been submitted elsewhere for a degree.

It is further certified that the student has fulfilled all requirements of Comprehensive Examination, Candidacy, and SOTA for the award of Ph.D. Degree.

Supervisor



Dr. Hari Prabhat Gupta

Associate Professor,

Department of Computer Science and Engineering,

Indian Institute of Technology (BHU) Varanasi,

Uttar Pradesh, INDIA 221005.

पर्यवेक्षक/Supervisor

संगणक विज्ञान एवं अभियांत्रिकी विभाग

Department of Computer Sc. & Engg

भारतीय प्रौद्योगिकी संस्थान

Indian Institute of Technology

(काशी हिन्दू विश्वविद्यालय)

(Banaras Hindu University)

वाराणसी/Varanasi-221005

COPYRIGHT TRANSFER CERTIFICATE

Title of the Thesis: Smart sensing-based approaches for Real-time River Water Pollution Monitoring

Name of the Student: Swati Sandeep Chopade

Copyright Transfer

The undersigned hereby assigns to the Institute of Technology (Banaras Hindu University) Varanasi all rights under copyright that may exist in and for the above thesis submitted for the award of the *Doctor of Philosophy*.

Date:

Place: Varanasi


(Swati Sandeep Chopade)

Note: However, the author may reproduce or authorize others to reproduce material extracted verbatim from the thesis or derivative of the thesis for author's personal use provided that the source and the Institute's copyright notice are indicated.

ACKNOWLEDGEMENT

First and foremost, I would like to thank my supervisor, Dr. Hari Prabhat Gupta, for his invaluable support and assistance. I feel immense pleasure in expressing my profound sense of gratitude and sincere regard for his constant feedback and expertise during all these years. I am eternally grateful to have had the opportunity to work on my thesis under his supervision.

I would like to thank the members of my Doctoral committee, Prof. Bhaskar Biswas, Department of Computer Science and Engineering, and Dr. Rajeev, Department of Mathematics, for their help and support throughout the tenure of my studies. Special thanks to Dr. Tanima Dutta for her consistent assistance in both work and life aspects. I would also like to convey my sincere gratitude to Dr. S. K. Singh, Head of Department of Computer Science and Engineering and all the RPEC and DPGC members for their suggestions and endorsement to this work.

I don't have words to thank my friend, brother, and mentor, Dr. Rahul Mishra, for his brilliant support and helping me to overcome the challenges I have faced in the development of this work. I am grateful to my colleagues and friends, Dr. Ashish, Dr. Preti Kumari, Ramakant Kumar, and Dr. Surbhi Saraswat for the long discussions and their brilliant insights that have helped me to overcome the challenges I have faced in the development of this work.

Finally, I express my heartfelt gratitude to my mother Mrs. Balika Kolekar, my husband Sandeep Chopade, my daughter Swarnili Chopade, and my son Swarnjit Chopade for their constant support, love, encouragement, and sacrifices. Their affectionate love and care cannot be expressed in words.

With limitless humility, I would like to praise and thank the “**Sankat Mochan Hanuman Ji**” and “**Baba Kashi Vishwanath Ji**”. The almighty, the Merciful compassionate who bestowed me with all the favourable circumstances to achieve the desired goal of life through this crucial juncture.

(Swati Sandeep Chopade)

List of Figures

1.1	Sensors used for deciding water quality.	4
1.2	Applications of water	5
2.1	Architecture and spreading factor of LoRa	19
2.2	Illustration of different entities of a game.	22
3.1	An example of river dataset collection and classification using deep learning based classifier.	34
3.2	Step-by-step procedure followed for river water data collection.	39
3.3	Steps involve in Lab dataset collection of the river water.	42
3.4	Variation in values of different parameters of lab data collected from Ganges in Varanasi.	44
3.5	Steps involve in sensory dataset collection of the river water.	45
3.6	Heat maps of different parameters in the sensory dataset [1], on a specific date (for river Ganges in Varanasi).	48
3.7	Portable waterproof Hanna multi-parameter HI-9829 meter	50
3.8	Illustration of first phase involve in the proposed approach.	51
3.9	Block diagram for estimation of water quality index.	52
3.10	Weight establishment in proposed approach using NSF [2].	57
3.11	Performance measure of river datasets during training.	62
3.12	Impact of the selected parameter(s) on proposed approach performance.	63
3.13	Class-wise accuracy. Part(a) for selected 6 parameters and Part(b) for all 17 parameters.	64
4.1	Block diagram for automatic annotation of sensory data.	72
4.2	Block diagram for construction of deep learning classifier II	77
4.3	Illustration of LSTM model to extract features from labeled sensory dataset to build classifier II	78
4.4	Illustration of a LSTM cell with different gates and cell state.	80

4.5	Illustration of all phases of sensor-based river water pollution assessment system.	82
4.6	Feature extraction of test instance followed by label prediction.	82
4.7	Performance measure for different combinations of LSTM layers and LSTM cells.	84
4.8	Value of precision and recall on training and testing sensory dataset of river water.	85
4.9	Illustration of noise handling capability of $\mathcal{L}(\cdot)$ with combination of variable parameters.	86
4.10	Class-wise accuracy of proposed approach on the data collected from different rivers (<i>i.e.</i> , Ganges, Godavari, Yamuna, and Hindon).	86
4.11	Accuracy of existing work on 0% and 20% noise concentration. (A1 : default loss and A2 : noise handling loss).	90
5.1	Illustration of a river water pollution monitoring using LoRa network in IoT.	94
5.2	Block diagram for the estimation of WQI.	99
5.3	Teacher model (\mathbf{M}^t) for estimating the WQI using combination of spatial, temporal, and hand-crafted features.	100
5.4	Framework for training student on reduced dataset \mathcal{D}' under the guidance of teacher trained on dataset \mathcal{D}	102
5.5	Heat maps of different parameters in the sensory dataset [1], on a specific date (for river Ganges in Varanasi).	106
5.6	Performance results of models \mathbf{M}^t , \mathbf{M}_1^s , \mathbf{M}_2^s and \mathbf{M}_3^s during training with and without Knowledge Distillation (KD). Parts (a1)-(a4) and (b1)-(b4) consider RWM dataset versions RWM ₁ and RWM ₂ having six ($\mathbf{a}_1 - \mathbf{a}_6$) and five ($\mathbf{a}_1 - \mathbf{a}_5$) classes, respectively.	109
5.7	Class-wise accuracy of teacher on dataset versions RWM ₁ and RWM ₂	110
5.8	Impact of hyper-parameters (μ_1 and μ_2) on student models.	111
5.9	Class-wise performance of student models on RWM ₁ and RWM ₂	111
6.1	Block diagram of energy-efficient pollution data communication system.	123
6.2	Block diagram of ERWM system.	132
6.3	Heat maps of different parameters in the sensory dataset [1], on a specific date (for river Ganges in Varanasi).	133
6.4	Illustration of the number of LNs on different game parameters, where TT is Transmission Time.	134

6.5	Illustration of energy, accuracy, and execution time on different schemes (\mathbf{S}_1 , \mathbf{S}_2 , and \mathbf{S}_3) for devices (\mathbf{d}_1 , \mathbf{d}_2 , \mathbf{d}_3 , and \mathbf{d}_4).	136
6.6	Illustration of devices used for sensing the river water.	138
6.7	Illustration of results of real world evaluation.	138

List of Tables

1.1	Example: Healthy Drinking Water	4
2.1	A comparative summary of the existing work for handling noisy labels in the dataset.	24
2.2	A Comparative summary of the existing literature on water pollution assessment. <i>#TE</i> , <i>#TU</i> , <i>#S</i> , <i>#A</i> , <i>#N</i> , <i>#C</i> , <i>#ORP</i> , <i>#SA</i> , AND <i>#FR</i> denote Temperature, Turbidity, Sulphate, Ammonia, Nitrate, Conductivity, Oxygen-Reduction Potential, Salinity, and Flow Rate, respectively.	28
2.3	A summary of existing work on water pollution monitoring.	30
3.1	Different Wireless Communication Protocols for water pollution monitoring	41
3.2	Lab dataset parameters along with its unit collected in our Lab dataset. <i>#NTU</i> , <i>#psi</i> , <i>#ppm</i> , <i>#cfu</i> , AND <i>#rfu</i> denote Nephelometric Turbidity unit, Pounds per square inch, parts per million or milligrams per liter (mg/L), Colony forming unit, and relative fluorescence unit, respectively	44
3.3	Sensory dataset parameters along with its unit collected in our sensory dataset. <i>#NTU</i> , <i>#psi</i> , <i>#ppm</i> , <i>#cfu</i> , <i>#μS/cm</i> ,AND <i>#rfu</i> denote Nephelometric Turbidity unit, Pounds per square inch, parts per million or milligrams per liter (mg/L), Colony forming unit, microSiemens per centimeter, and relative fluorescence unit, respectively	48
3.4	Range of WQI for different class labels and their corresponding integer representation used in this paper.	61
4.1	Precision (P), Recall (R), and F1-score (F1) using default and noise handling functions with different noise concentration.	88
4.2	Performance comparison of the proposed approach with the existing work.	90