Algorithms for Intelligent Systems
Series Editors: Jagdish Chand Bansal · Kusum Deep · Atulya K. Nagar

Rajendra Prasad Yadav
Satyasai Jagannath Nanda
Prashant Singh Rana
Meng-Hiot Lim *Editors*

Proceedings of the International Conference on Paradigms of Computing, Communication and Data Sciences

PCCDS 2022

Springer

Algorithms for Intelligent Systems

This book series publishes research on the analysis and development of algorithms for intelligent systems with their applications to various real world problems. It covers research related to autonomous agents, multi-agent systems, behavioral modeling, reinforcement learning, game theory, mechanism design, machine learning, meta-heuristic search, optimization, planning and scheduling, artificial neural networks, evolutionary computation, swarm intelligence and other algorithms for intelligent systems.

The book series includes recent advancements, modification and applications of the artificial neural networks, evolutionary computation, swarm intelligence, artificial immune systems, fuzzy system, autonomous and multi agent systems, machine learning and other intelligent systems related areas. The material will be beneficial for the graduate students, post-graduate students as well as the researchers who want a broader view of advances in algorithms for intelligent systems. The contents will also be useful to the researchers from other fields who have no knowledge of the power of intelligent systems, e.g. the researchers in the field of bioinformatics, biochemists, mechanical and chemical engineers, economists, musicians and medical practitioners.

The series publishes monographs, edited volumes, advanced textbooks and selected proceedings.

Indexed by zbMATH.

All books published in the series are submitted for consideration in Web of Science.

Rajendra Prasad Yadav ·
Satyasai Jagannath Nanda · Prashant Singh Rana ·
Meng-Hiot Lim
Editors

Proceedings of the International Conference on Paradigms of Computing, Communication and Data Sciences

PCCDS 2022

Springer

Editors
Rajendra Prasad Yadav
Department of Electronics
and Communication Engineering
Malaviya National Institute of Technology
Jaipur, Rajasthan, India

Satyasai Jagannath Nanda
Department of Electronics
and Communication Engineering
Malaviya National Institute of Technology
Jaipur, Rajasthan, India

Prashant Singh Rana
Department of Computer Science
and Engineering
Thapar Institute of Engineering
and Technology
Patiala, Punjab, India

Meng-Hiot Lim
School of Electrical and Electronic
Engineering
Nanyang Technological University
Singapore, Singapore

ISSN 2524-7565 ISSN 2524-7573 (electronic)
Algorithms for Intelligent Systems
ISBN 978-981-19-8741-0 ISBN 978-981-19-8742-7 (eBook)
https://doi.org/10.1007/978-981-19-8742-7

Preface

This book contains outstanding research papers as the proceedings of the International Conference on Paradigms of Communication, Computing and Data Sciences (PCCDS 2022). PCCDS 2022 has been organized by Malaviya National Institute of Technology Jaipur, India, and technically sponsored by Soft Computing Research Society, India. The conference is conceived as a platform for disseminating and exchanging ideas, concepts, and results of researchers from academia and industry to develop a comprehensive understanding of the challenges of the advancements of intelligence in computational viewpoints. This book will help in strengthening congenial networking between academia and industry. We have tried our best to enrich the quality of the PCCDS 2022 through the stringent and careful peer-review process. This book presents novel contributions to Communication, Computing and Data Sciences and serves as reference material for advanced research.

PCCDS 2022 received many technical contributed articles from distinguished participants from home and abroad. PCCDS 2022 received 349 research submissions from 20 different countries, viz. Bangladesh, China, Germany, Greece, Iceland, India, Indonesia, Malaysia, Mexico, Morocco, Philippines, Poland, Qatar, Romania, Russia, Senegal, Serbia, Spain, Ukraine, and the USA. After a very stringent peer-reviewing process, only 62 high-quality papers were finally accepted for presentation and the final proceedings.

Jaipur, India — Rajendra Prasad Yadav
Jaipur, India — Satyasai Jagannath Nanda
Patiala, India — Prashant Singh Rana
Singapore — Meng-Hiot Lim

Contents

About the Editors

Prof. Rajendra Prasad Yadav is currently working as a Professor-HAG in the Department of Electronics and Communication Engineering, Malaviya National Institute of Technology, Jaipur, Rajasthan, India. He has more than four decades of teaching and research experience. He was instrumental in starting new B.Tech, M.Tech courses and formulating Ph.D Ordinances for starting research work in Rajasthan Technical University (RTU) Kota and other affiliated Engg colleges as Vice Chancellor of the University. He has served as HOD of Electronics and Comm. Engg., President Sports and Library, Hostel warden, Dean Student Affairs at MNIT Jaipur. At present he is also the Chief Vigilance Officer of MNIT Jaipur since 2015. Prof. Yadav received the Ph.D degree from MREC Jaipur and M.Tech degree from IIT Delhi. Under his supervision 15 Ph.D students have received Ph.D degree, 7 students are working for their Ph.D degree. Forty M.Tech students have carried out their dissertation work under his guidance. He has published more than 200 peer review research papers which has received 1150 citations. His research interest are Error Control Codes and MIMO-OFDM, RF and Antenna Systems, Mobile and Wireless Communication Systems, Optical Switching and Materials, Mobile Adhoc and Sensor Network, Device Characterization and MEMS, Cognitive Radio Networks.

Dr. Satyasai Jagannath Nanda is an assistant professor in the Department of Electronics and Communication Engineering, Malaviya National Institute of Technology Jaipur since June 2013. Prior to joining MNIT Jaipur he has received the PhD degree from School of Electrical Sciences, IIT Bhubaneswar and M. Tech. degree from Dept. of Electronics and Communication Engg., NIT Rourkela. He was the recipient of Canadian Research Fellowship- GSEP, from Dept. of Foreign Affairs and Intern. Trade (DFAIT), Govt. of Canada for the year 2009-10. He was awarded Best PhD thesis award at SocPros 2015 by IIT Roorkee. He received the best research paper awards at SocPros-2020 at IIT Indore, IC3-2018 at SMIT Sikkim, SocPros-2017 at IIT Bhubaneswar, IEEE UPCON-2016 at IIT BHU and Springer OWT-2017 at MNIT. He is the recipient of prestigious IEI Young Engineers Award by Institution of Engineers, Govt. of India in the field of Electronics and Telecommunication Engineering for the year 2018-19. Dr. Nanda is a Senior Member of IEEE and IEEE

Computational Intelligence Society. He has received travel and research grants from SERB, UGC, CCSTDS (INSA), INAE. Till date he has published 40 SCI/SCOUPUS Journal articles and 50 international conference proceedings which received almost twelve hundred citations. He is the in-charge of Digital Signal and Image Processing (DSIP) Lab. at MNIT Jaipur. Under his supervision at MNIT Jaipur six researchers have awarded PhD and five researchers are continuing their research work. Along with it he has supervised 22 M. Tech thesis. Dr. Nanda is co-coordinator of Electronics and ICT Academy at MNIT Jaipur which is a set-up of Ministry of Electronics and IT, Govt. of India of Grant 10 Crore.

Dr Prashant Singh Rana is presently working as Associate Professor in the Computer Science and Engineering Department, Thapar Institute of Engineering & Technology, Patiala, Punjab. He received his both PhD and MTech from ABV-IIITM, Gwalior. His areas of research are Machine Learning, Deep Learning, Bioinformatics and Optimization. He has published more than 70 research papers in different journals and conferences. He completed five projects sponsored by DST, ICMR, NVIDIA and one projects are going on. He published 10 patents. He guided seven PhD students and 18 Masters students.

Dr Meng-Hiot Lim is a faculty in the School of Electrical and Electronic Engineering. He is holding a concurrent appointment as a Deputy Director for the M.Sc in Financial Engineering and the Centre for Financial Engineering, anchored at the Nanyang Business School. He is a versatile researcher with diverse interests, with research focus in the areas of computational intelligence, evolvable hardware, finance, algorithms for UAVs and memetic computing. He is currently the Editor-in-Chief of the Journal of Memetic Computing published by Springer. He is also the Series Editor of the book series by Springer titled “Studies in Evolutionary Learning and Optimization”.

Chapter 1
Optimized Watermarking Scheme for Copyright Protection of Medical Images

Rohit Thanki and Purva Joshi

1 Introduction

Image sharing is easy with today's open-access media. However, attackers or imposters can manipulate the images on open-access media. This leads to copyright issues. Because of this, if an image is shared on an open-access medium, it must be protected by copyright. Watermarking can be used to solve this problem [1–11]. Using an embedding factor, watermarked content is generated from the cover image. There are many types of watermarking based on the content of their covers, the processing domains, the attack, and the extraction method [1, 6]. Watermarks can be applied as text watermarks, image watermarks, or signal watermarks, depending on the content of the cover. Watermarking can be classified into four types based on the processing domain: spatial domain, transform domain, a hybrid domain, and sparse domain. According to their resistance to attacks, watermarks can be classified as robust or fragile. Additionally, blind, non-blind, and semi-blind watermarking can be classified as types of watermarking based on how they are extracted.

According to the literature [1–11], watermark embedding is performed using an embedding factor. The watermark information is inserted or modified into the cover medium's content using the embedding factor value. In the literature, researchers have created watermarked information using their embedding factors. Unfortunately, there is no standard for embedding factors. Therefore, an optimized process is required for embedding factor standardization. The optimization process will identify the optimal embedding factor to produce the best results for the watermarking algorithm.

R. Thanki (✉)
Senior Member, IEEE Gujarat Section, Rajkot, India
e-mail: rohitthanki9@gmail.com

P. Joshi
University of Pisa, Pisa, Italy

R. P. Yadav et al. (eds.), *Proceedings of the International Conference on Paradigms of Computing, Communication and Data Sciences*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-19-8742-7_1

The watermarking algorithm uses various optimization algorithms to find an optimal embedding factor. The embedding factor is determined by watermarking evolution parameters (called a fitness function, f). Watermark transparency is measured by peak signal-to-noise ratios (PSNR) or weighted PSNRs. Structural similarity index measures (SSIMs), bit error rates (BERs), and bit correction rates (BCRs) are used to measure robustness. Watermark size and cover data size are used to calculate payload capacity. Transparency, robustness, and payload capacity are measured by PSNR, NC, and PC, respectively. As a result, the fitness function may look like this:

$$f_m = \mathrm{PSNR}_m + w_1 \cdot \mathrm{NC}_m + w_2 \cdot \mathrm{PC}_m \tag{1}$$

where m is a no. of iteration and w_1 and w_2 are weighted factors.

The embedding factors are used to generate an image with a watermark. To obtain a watermarked image, embedding factors were used according to the researchers' specifications. It is time consuming to select embedding factors manually. The best embedding factor is thus determined by algorithms regarding watermarking techniques. Embedding factors are affected by cover data, watermarks, and embedding processes. An array of optimization algorithms can currently be used to accomplish this, including genetic algorithms (GAs) [12], particle swarm optimizations (PSOs) [13], genetic programming (GP) [14, 15], differential evaluations (DAs) [16], simulated annealing (SAs) [17], tabu searches [18], colony optimizations [19], and harmony searches [20]. Due to its ease of understanding and implementation, the PSO algorithm is widely used for optimizing embedding factors in watermarking [21].

The literature contains a few watermarking schemes using PSO to protect medical images from copyright infringements [22–25]. Findik et al. [22] proposed a PSO watermarking scheme for cover color images. A pseudo-random noise generator was used to determine where the watermark image should be embedded. The watermark bit was inserted into a block of the cover image using PSO after finding the best pixel locations in this block. The payload capacity of this scheme is smaller than that of a blind scheme. Watermarking schemes based on PSO and discrete wavelet transform (DWT) have been proposed by Fakhari et al. [23] and Wang et al. [24]. An image with a grayscale medical cover and a greyscale standard was proposed by the Fakhari scheme, while the Wang scheme proposed the standard grayscale image. It was a non-blind scheme in both cases. There is, however, a limited payload capacity with these two schemes. Wang's scheme is robust against standard watermarking attacks, whereas Fakhari's is robust against geometric attacks. A DWT watermarking scheme was presented by Chakraborty et al. [25]. A detailed wavelet sub-band of the host image is augmented with PN sequences by watermark bits. The embedding factors for watermarks are found using PSO. The authors do not discuss the scheme's robustness against attacks in their paper.

A new blind watermarking scheme is developed and proposed for medical images in this paper to overcome some limitations of the existing techniques [22, 23]. Watermarking schemes using RDWTs offer better transparency and payload capacity than

those using wavelet transforms alone. This scheme has the following main features: (1) uses RDWT and PSO properties to overcome some of the limitations of standard watermarking procedures, such as selecting the embedding factor and limited payload capacity; (2) watermark images can be blindly extracted, which cannot be done using the existing scheme [23, 25]. The proposed scheme is robust based on the results. Moreover, the proposed scheme provides greater transparency and payload capacity than the existing schemes such as Fakhari's scheme [23] and Chakraborty's scheme [25]. The proposed scheme selects embedding factors according to an optimization mechanism. Watermarking schemes are improved by adding a mechanism for optimization to the traditional trade-offs.

It proceeds as follows: Sect. 2 describes a proposed blind and optimized watermarking scheme. A comparison of the proposed scheme and the experimental results is presented in Sect. 3. The conclusion of the paper is presented in Sect. 4.

2 Proposed Scheme

The PSO algorithm implements a robust and blind watermarking scheme [25]. This scheme uses digital wavelet transform (DWT). However, the payload capacity of this scheme is very low. Additionally, watermarked images are less transparent with this scheme. Moreover, this scheme applies only to certain types of images or signals. Using block RDWT and PSO, we propose an approach to embed the monochrome watermark directly into the LH, HL, and HH wavelet sub-bands of the cover medical image. HH, HL, and LH, which contain detailed wavelet sub-bands, provide less visual information for embedding watermarks. As a result, the proposed scheme embeds the watermark in the detailed wavelet sub-bands. Sub-bands are divided into non-overlapping blocks. Watermark bits 0 and 1 are embedded using two uncorrected noise sequences. As a result of each noise sequence, the coefficients of its corresponding sub-band are modified using the optimal embedding factor. The PSO algorithm determined an optimal embedding factor. Embedding and extraction are two processes in the proposed scheme.

2.1 Embedding Process

The steps for the embedding process are given below.

Step 1. A single-level RDWT is used to separate the cover medical image into wavelet sub-bands such as LL (approximation sub-band) and LH, HL, and HH (detail sub-bands).
Step 2. Monochrome watermark images are converted into binary sequences.
Step 3. Create non-overlapping blocks from detailed wavelet sub-bands LH, HL, and HH.

Step 4. The block size is equal to the size of two uncorrelated noise sequences generated by a noise generator. S_0 and S_1 are the noise sequences for watermark bits 0 and 1.
Step 5. The watermark sequence modifies the coefficients of detailed wavelet sub-bands based on an optimal embedding factor (k) which is obtained from PSO. For each block of cover medical image, this procedure was conducted for all coefficients.
Step 6. A modified wavelet sub-band is the output as the result of step 5. After that, apply single-level inverse RDWT to these modified sub-bands and unmodified approximation sub-bands in order to generate the watermarked medical image.

2.2 Extraction Process

The steps for the extraction process are given below.

Step 1. A watermarked medical image is decomposed into different wavelet sub-bands using a single-level RDWT, including approximation sub-bands and detailed sub-bands like WLH, WHL, and WHH. WLH, WHL, and WHH should be converted into non-overlapping blocks.
Step 2. Take uncorrelated noise sequences that are generated during embedding.
Step 3. Using the correlation between the noise sequences (S_0, S_1) and the detailed wavelet coefficients (WLH, WHL, and WHH), recover the watermark bit from the detailed wavelet coefficients of the watermarked medical image. The correlation result of coefficients with noise sequence S_1 and noise sequence S_0 indicates C_1 and C_0.
Step 4. Whenever $C_0 > C_1$, bit 0 is selected as the watermark bit value. Otherwise, bit 1 is chosen.
Step 5. The recovered watermark image is created by reshaping the extracted sequence into the matrix.

2.3 Generation of Optimal Embedding Factor

Any watermarking scheme depends on the embedding factor to meet its basic requirements. Watermarked images with large embedding factors have lower transparency, but recovered watermarked images with high embedding factors have better quality. Many existing schemes in the literature keep the embedding factor constant, but these do not work well with multimedia data. As a result, some adaptive schemes are needed to calculate the appropriate embedding factors for various multimedia data types. This paper combines a block RDWT-based watermarking scheme with a popular optimization algorithm, PSO, to provide an optimal embedding factor value. According to the proposed scheme, PSNR and NC can be used to calculate fitness

Table 1 Obtained optimal embedding factor for proposed scheme

Embedding factor range	k_1	k_2	k_3
0.0–1.0	0.5483	0.6346	0.8362
0.0–2.0	1.5718	1.8092	1.6667
0.0–3.0	2.0803	2.7933	2.6394
0.0–4.0	3.8015	3.8167	3.8451
0.0–8.0	7.6857	6.2683	5.6797
0.0–10.0	8.0542	9.8810	7.6335
0.0–50.0	42.8845	49.0747	44.0159
0.0–150.0	143.4951	148.4743	142.2608
0.0–250.0	233.1310	225.7728	218.1750

functions for each particle's population. Optimal solutions (gbest) are selected by selecting the maximum fitness values. Equation (2) gives the fitness function:

$$\text{fitness} = \text{PSNR}(C, \text{WA}) + \left(100 * \text{NC}\left(w, w'\right)\right) \tag{2}$$

The above equation represents the peak signal-to-noise ratio defined by PSNR and the normalized correlation by NC. A cover medical image is indicated by variable C, while variable WA indicates a watermarked medical image, and a watermark image is indicated by variables w, w'. According to the experimental results, the proposed fitness function works well. In this case, the PSO parameters are selected to help compare schemes, their values are provided as constants $C1$ and $C2$, the number of particles is 5, the number of iterations is 8, and the initial weight α is 0.9. Some trial and experimental medical images found the best range between 0.0 and 250.0. The PSO algorithm provided the optimal embedding factor given in Table 1. The embedding factor k is represented in Table 1 for the generation of watermarked medical images.

3 Experimental Results

An MRI image with grayscale (512 × 512 pixels) [26] is used as a cover medical image and the hospital logo (64 × 64) as a watermark image in testing the proposed scheme (Fig. 1). In this example, the watermark image is inserted directly into the cover medical image. The resultant images using the proposed scheme are shown in Fig. 2.

Table 2 summarizes the PSNR and NC values based on the proposed scheme. Using this measurement, the PSNR measures the imperceptibility (transparency) of the embedded watermark in the cover image. As part of the study, NC also measured the robustness of extracting watermarks from watermarked medical images. A T_{EMB} (s) and T_{EXT} (s) indicate how long it takes to embed the watermark into a cover

Fig. 1 **a** Cover (MRI) image, **b** watermark image

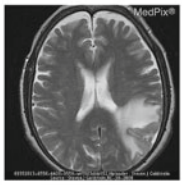
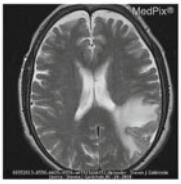
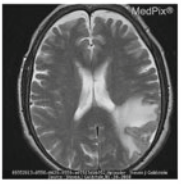
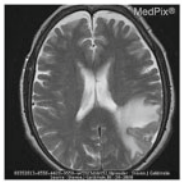
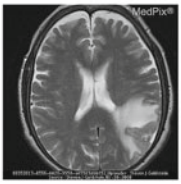
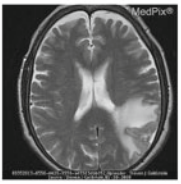
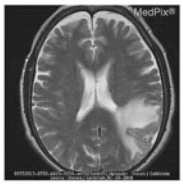
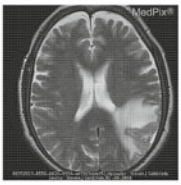
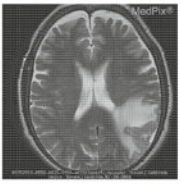
Embedding Factor Range	**0.0 – 1.0**	**0.0 – 2.0**	**0.0 – 3.0**
Watermarked Image			
Recovered Watermark Image			
Embedding Factor Range	**0.0 – 4.0**	**0.0 – 8.0**	**0.0 – 10.0**
Watermarked Image			
Recovered Watermark Image			
Embedding Factor Range	**0.0 – 50.0**	**0.0 – 15.0**	**0.0 – 250.0**
Watermarked Image			
Recovered Watermark Image			

Fig. 2 Generated watermarked medical images and recovered watermark image

medical image and how long it takes to extract the watermark from that image. In total, this algorithm generates watermarked medical images in around 3 s.

Table 3 summarizes the NC values of recovered watermark images for different watermarking attacks. This scheme can also provide robustness for medical images, as shown by the results. This indicates that telemedicine applications can use this scheme to secure medical images. Based on Table 4, the proposed scheme is compared with the Fakhari and Chakraborty schemes that provide similar copyright protection for medical images.

For embedding the watermark image, the Fakhari scheme [23] and the Chakraborty scheme [25] use DWT, whereas the proposed scheme uses RDWT. Fakhari's scheme [23] has a payload capacity of 10 bits, while Chakraborty's scheme

Table 2 Quality measures values of proposed scheme

Range of embedding factor	PSNR (dB)	NC	T_{EMB} (s)	T_{EXT} (s)
0.0–1.0	63.02	0.6805	1.9943	1.5326
0.0–2.0	55.05	0.7872	1.4002	1.5577
0.0–3.0	51.58	0.8493	1.4411	1.5326
0.0–4.0	47.95	0.9035	1.4463	1.5061
0.0–8.0	43.14	0.9475	1.3609	1.6488
0.0–10.0	40.88	0.9667	1.3595	1.5823
0.0–50.0	26.44	0.9862	1.5518	1.7159
0.0–150.0	16.37	0.9975	1.5061	1.6962
0.0–250.0	12.50	0.9996	1.5345	1.6310

Table 3 NC values of proposed scheme under different watermarking attacks

Attacks	NC	Attacks	NC
JPEG ($Q = 80$)	0.9996	Motion blurring	0.8067
JPEG ($Q = 25$)	0.9993	Gaussian blurring	0.9996
Median filtering ($3 \cdot 3$)	0.9979	Sharping	1.0000
Gaussian noise (variance = 0.005)	0.9996	Histogram equalization	1.0000
Salt and pepper noise (variance = 0.005)	0.9996	Rotation (20°)	0.5649
Speckle noise (variance = 0.005)	0.9996	Cropping (20%)	0.9996
Intensity adjustment	1.0000	Scaling (512–256–512)	0.9755

Table 4 Performance comparison of proposed scheme with the existing schemes [23, 25]

Features	Fakhari scheme [23]	Chakraborty scheme [25]	Proposed scheme
Used transform	DWT	DWT	RDWT
$PSNR_{Max}$ (dB)	51.55	25.7253	63.02
Payload capacity (bits)	10	1024	4096

Table 5 Performance comparison of proposed scheme with recently published schemes (2022, 2021) [23, 25]

Features	Rezaee scheme [27]	Sharma scheme [28]	Golda scheme [29]	Proposed scheme
Used optimization algorithm	Whale	Firefly	Social group	Particle swarm optimization
$PSNR_{Max}$ (dB)	39.87	57.58	23.78	63.02
NC_{Max}	0.9807	Not reported	Not reported	0.9996

[25] has a payload capacity of 1024 bits, which is less than the proposed scheme's payload capacity. Compared to the Fakhari scheme [23] and the Chakraborty scheme [25], the proposed scheme performs better in transparency and payload capacity. Furthermore, with higher PSNR values, the transparency of watermark medical images is improved. As a result, Table 4 shows that the proposed scheme has a higher PSNR than the existing schemes, which implies that it has a higher degree of transparency.

Table 5 summarizes the performance comparison of the proposed scheme with that of recently published schemes (2022, 2021) [27–29]. The comparison is based on PSNR and NC values and an optimization algorithm. Compared with recently published schemes [27–29], this proposed scheme outperformed them.

4 Conclusion

For copyright protection of medical images, we present a watermarking scheme based on RDWT and PSO. In this case, the RDWT is used to increase payload capacity, while the PSO is used to generate optimal embedding factors. The proposed scheme for embedding a secret logo into medical images for copyright protection was secure and accurate. However, as a limitation of this proposed scheme, it can only watermark binary watermark images. Additionally, the proposed scheme performed better than the existing schemes.

References

1. Langelaar GC, Setyawan I, Lagendijk RL (2000) Watermarking digital image and video data. A state-of-the-art overview. IEEE Signal Process Mag 17(5):20–46
2. Thanki R, Borra S, Dwivedi V, Borisagar K (2017) An efficient medical image watermarking scheme based on FDCuT–DCT. Eng Sci Technol Int J 20(4):1366–1379
3. Lakshmi HR, Surekha B, Raju SV (2017) Real-time implementation of reversible watermarking. In: Intelligent techniques in signal processing for multimedia security. Springer, Cham, pp 113–132

4. Thanki R, Borra S (2018) A color image steganography in hybrid FRT–DWT domain. J Inf Secur Appl 40:92–102
5. Thanki R, Dwivedi V, Borisagar K, Borra S (2017) A watermarking algorithm for multiple watermarks protection using SVD and compressive sensing. Informatica 41(4):479–493
6. Borra S, Lakshmi H, Dey N, Ashour A, Shi F (2017) Digital image watermarking tools: state-of-the-art. In: Information technology and intelligent transportation systems: proceedings of the 2nd international conference on information technology and intelligent transportation systems, vol 296, Xi'an, China, p 450
7. Surekha B, Swamy GN (2013) Sensitive digital image watermarking for copyright protection. IJ Netw Secur 15(2):113–121
8. Surekha B, Swamy G, Reddy KRL (2012, July) A novel copyright protection scheme based on visual secret sharing. In: 2012 third international conference on computing communication & networking technologies (ICCCNT). IEEE, pp 1–5
9. Dey N, Roy AB, Das A, Chaudhuri SS (2012, October) Stationary wavelet transformation based self-recovery of blind-watermark from electrocardiogram signal in wireless telecardiology. In: International conference on security in computer networks and distributed systems. Springer, Berlin, Heidelberg, pp 347–357
10. Dey N, Dey G, Chakraborty S, Chaudhuri SS (2014) Feature analysis of blind watermarked electromyogram signal in wireless telemonitoring. In: Concepts and trends in healthcare information systems. Springer, Cham, pp 205–229
11. Dey N, Ashour AS, Chakraborty S, Banerjee S, Gospodinova E, Gospodinov M, Hassanien AE (2017) Watermarking in biomedical signal processing. In: Intelligent techniques in signal processing for multimedia security. Springer, Cham, pp 345–369
12. Holland JH (1975) Adaptation in natural and artificial systems. University of Michigan Press
13. Kennedy J, Eberhart RC (1995) Particle swarm optimization. In: Proceedings of the 1995 IEEE international conference on neural networks, Perth, Australia, pp 1942–1948
14. Koza JR (1992) Genetic programming: on the programming of computers by means of natural selection. MIT Press
15. Koza JR (1994) Genetic programming as a means for programming computers by natural selection. Stat Comput 4(2):87–112
16. Storn R, Price K (1997) Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. J Global Optim 11(4):341–359
17. Kirkpatrick S, Gelatt CD, Vecchi MP (1983) Optimization by simulated annealing. Science 220(4598):671–680
18. Glover F (1977) Heuristics for integer programming using surrogate constraints. Dec Sci 8(1):156–166
19. Dorigo M, Maniezzo V, Colorni A (1996) Ant system: optimization by a colony of cooperating agents. IEEE Trans Syst Man Cybern Part B 26(1):29–41
20. Geem ZW, Kim JH, Loganathan GV (2001) A new heuristic optimization algorithm: harmony search. Simulation 76(2):60–68
21. Li X, Wang J (2007) A steganographic method based upon JPEG and particle swarm optimization algorithm. Inf Sci 177(15):3099–3109
22. Findik O, Babaoğlu İ, Ülker E (2010) A color image watermarking scheme based on hybrid classification method: particle swarm optimization and k-nearest neighbor algorithm. Opt Commun 283(24):4916–4922
23. Fakhari P, Vahedi E, Lucas C (2011) Protecting patient privacy from unauthorized release of medical images using a bio-inspired wavelet-based watermarking approach. Digital Signal Process 21(3):433–446
24. Wang YR, Lin WH, Yang L (2011) An intelligent watermarking method based on particle swarm optimization. Expert Syst Appl 38(7):8024–8029
25. Chakraborty S, Samanta S, Biswas D, Dey N, Chaudhuri SS (2013, December) Particle swarm optimization-based parameter optimization technique in medical information hiding. In: 2013 IEEE international conference on computational intelligence and computing research (ICCIC), IEEE, pp 1–6

26. MedPixTM Medical Image Database available at http://rad.usuhs.mil/medpix/medpix.html, https://medpix.nlm.nih.gov/home. Last access year: 2021
27. Rezaee K, SaberiAnari M, Khosravi MR (2022) A wavelet-based robust medical image watermarking technique using whale optimization algorithm for data exchange through internet of medical things. In: Intelligent healthcare. Springer, Singapore, pp 373–394
28. Sharma S, Choudhary S, Sharma VK, Goyal A, Balihar MM (2022) Image watermarking in frequency domain using Hu's invariant moments and firefly algorithm. Int J Image Graph Signal Process 2:1–15
29. Golda D, Prabha B, Murali K, Prasuna K, Vatsav SS, Adepu S (2021) Robust image watermarking using the social group optimization algorithm. Mater Today Proc

Chapter 2
MobileNet + SSD: Lightweight Network for Real-Time Detection of Basketball Player

Banoth Thulasya Naik and Mohammad Farukh Hashmi

1 Introduction

In the field of computer vision, sports video analysis is one of the important topics. Many kinds of research, in particular, emphasize field sports like basketball, soccer, and field hockey, which are extremely popular outdoor sports all over the world. Analysis of field sports videos can be used for a variety of purposes, including event detection and player/team activity analysis. Low-level structural processes, such as player detection, classification, and tracking, are required for high-level applications. Player detection is a challenging issue though it is generally the first step in sports video analysis. The reason is basketball being an extremely dynamic sport, with players continuously changing their positions and postures. This paper offers a robust and efficient system for real-time player detection in basketball. This system enables detection of the position of the player in sports videos. As a result, only, one camera is displayed at any given moment. The player's size does not always remain consistent when the camera moves from one side of the court to other and zooms in at various points. Furthermore, some body parts may appear to be foreshortened in comparison to others as a result of players changing their angle in relation to the camera. Moreover, it is extremely common for a player to be partially covered by other players. This paper focuses on basketball player detection using a dynamic camera infrastructure.

In field sports, player detection algorithms must deal with a wide range of challenges, including changing lighting and weather conditions, as well as positional

B. T. Naik (✉) · M. F. Hashmi
Department of Electronics and Communication Engineering, National Institute of Technology, Warangal, India
e-mail: thulasyramsingh@student.nitw.ac.in

M. F. Hashmi
e-mail: mdfarukh@nitw.ac.in

R. P. Yadav et al. (eds.), *Proceedings of the International Conference on Paradigms of Computing, Communication and Data Sciences*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-19-8742-7_2

changes of players in pictures, such as size and rotation, depending on the camera viewpoint. Depending on the distance from the camera and the direction in which they travel, players may appear at varying sizes, resolutions, and orientations. As there is such a broad range of player uniform colors and textures, the team uniform and lighting have a significant impact on player appearance. An approach is proposed to address these problems for real-time player detection in basketball sports.

The rest of the paper is organized as follows. Section 2 provides a review of the existing literature related to player detection. The proposed methodology which is used to achieve the desired results is discussed in Sect. 3. The experimental results and performance metrics related to the work are presented in Sect. 4. Finally, the conclusion and scope of future work is discussed in Sect. 5.

2 Literature Survey

One of the most basic tasks in computer vision is extracting information from images. Researchers have developed systems that use structure-from-motion [1] to obtain geometric information for detection and classification to find semantic information. For a wide range of applications, detecting players from images and videos is significant [2]. Intelligent broadcast systems, for example, employ player positions to influence broadcast camera viewpoints. For instance, broadcast systems employ player positions to influence broadcast camera viewpoints [3, 4]. In [5] proposed a mechanism to classify group activities by detecting players. Player detection and team categorization also offer metadata for player tracking [6], player pose estimate, and team strategy analysis [7]. Player detection has attracted a lot of attention as a subset of object detection in sports. [8–10] presented background subtraction-based approaches to achieve real-time response in a basketball game. Nonetheless, in order to detect foreground objects reliably, all techniques assume the camera is stationary. Several learning-based approaches, such as Faster R-CNN [11] and YOLO [12], can be modified to identify players with high detection accuracy; however, due to poor pixel resolution, they may miss distant players.

3 Methodology

Detection algorithm comprises dual mechanism, i.e., a backbone and head. As essence, the backbone is often a pre-trained network for classifying the images. Here, the network called MobileNet [13] is used as backbone which is trained using over a million images while SSD [14] is used as head as shown in Fig. 1. Therefore, SSD contains various fixed layers, and results were defined in terms of classes of predicted and ground truth bounding boxes at the final 1-dimensional fully connected layer.

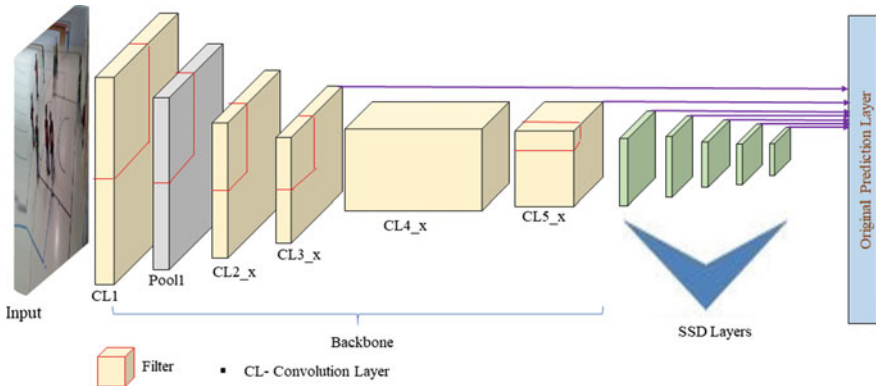


Fig. 1 Architecture of MobileNetv1 + SSD detector

3.1 SSD Parameters

3.1.1 Grid Cell and Anchor Box

Player detection in image entails determining the class and position of an element in the immediate vicinity. For example, an image of 4 × 4 grid is shown in Fig. 2. Here, the grating concentrates on producing appropriate spacing and form. The anchor boxes are accessible fields that are provided to complete that portion of the picture which has distinct parts in the grating separately.

Individual grid cells were assigned to SSD with various anchors or prefixes. The anchor boxes have complete control over the shape and size of each grid cell. Figure 2

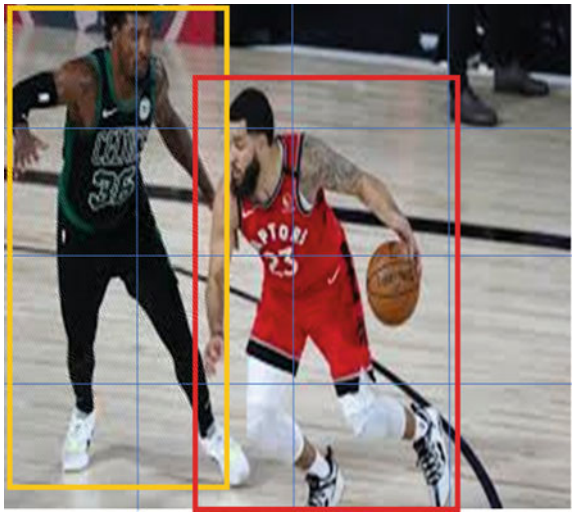


Fig. 2 Example of 4 × 4 grid and different size anchor boxes

shows two players, one with a height of anchor box and the other with a width of anchor box, indicating that the anchor boxes are of various sizes. The class and location of an item is finalized by anchor boxes with a lot of intersections through it. This information is utilized to train the network as well as to forecast the detected object's position once the network has been completed.

3.1.2 Zoom Level and Aspect Ratio

The size of the anchor boxes does not have to be the same as the grid cells. It is used to determine the degree to which the anchor box wants the posterior to move upward or downward using grid cells. As illustrated in Fig. 2 based on varying grades, certain items in form are broader while others are longer. The SSD architecture allows anchor boxes to have a higher aspect ratio. The various aspect ratios of anchor box links are described using the range of proportions.

3.1.3 Receptive Field

The permissible field input area is divided into a viewable zone using a specific convolutional neural network. Zeiler et al. [15] presented them as a back-line combination at the relative location using a distinguishing attribute and an actuation. The characteristics of distinct layers indicate the varying sizes of regions in the image due to the hampered process. This is done by pacing the bottom layer (5×5) first, followed by a console, resulting in a middle layer (3×3) with a single green pixel representing a 3×3 section of the input layer (bottom layer), as seen in Fig. 3. The convolution is then applied to the middle layer (green), with the upper red layer (2×2) containing the individual attribute equaling the input image's 7×7 region. The feature map is a green and red 2D array that points to a collection of features produced using a comparable feature extractor at different places of the input map in the form of an indication window. A comparable field exists for similar map features, which attempts to find similar patterns in different locations. As a result, a convolutional network at the local level is constructed.

3.2 Implementation of Proposed Methodology

TensorFlow, Keras, and OpenCV libraries were used to build deep learning models for player detection, which is akin to real-time object detection. First, the system was trained with known data, i.e., labeled basketball data so that a player who comes in any unseen frames or video can be detected. This paper deals with a pre-trained lightweight player detection model which is trained on third-party objects, and most of the objects are included in the training class except player. So, some of the layers of the proposed method were modified to train the model on labeled basketball

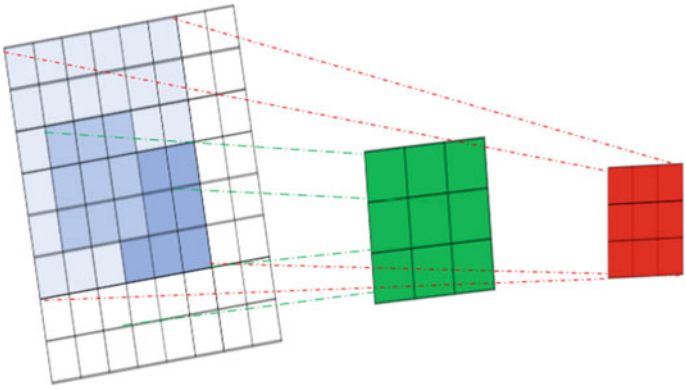


Fig. 3 Feature map visualization and the receptive field

data. Finally, combining the pre-trained network and the SSD approach, the system was ready to detect basketball players. Here, the pre-trained model (backbone) is MobieNetv1 which was combined with SSD to complete the architecture of player detection. Now, the architecture was trained by detecting labels from the training dataset based on the bounding boxes. Frames must be considered fixed resolution (512 × 512) for training the player detection model.

MobileNetv1 was utilized as the backbone of SSD architecture to improve the detection accuracy and frame rate. To detect multiple players, this technique must take a single shot. So, SSD is the best architecture designed based on neural networks to detect objects in a frame or video. Other techniques such as R-CNN, Fast R-CNN, and Faster R-CNN require two shots. This increases the computation time as parameters increase, which reduces speed of detection. The SSD method divides the bounding output space into a sequence of default boxes with different proportions and sizes and recognizes the restricted output space as the default box. The network quickly analyzes a default box for the existence of particular object classes and combines the boxes to detect an exact object. This network also encompasses a variety of models of varying sizes of natural adhesives and resolutions. In case if no object is present in the frame, then it is considered as background and ignored.

4 Dataset and Experimental Results

The proposed model carried out training on the basketball dataset [16] which was filmed during a basketball match (Resolution 1020 × 1080), and it contains a variable number of players, i.e., some frames contain 10 or 11 players while other frames contain 12 or 13 players of total 50,127 frames of which 40,300 frames are for training and 9827 frames are for testing. While training the proposed model, the resolution of frames was modified to 512 × 512, and various data augmentation techniques

Table 1 Configurations of experimental setup

Model training/testing setup	
Names	Experimental configuration
OS	Windows 10 Pro
CPU/GHz	Intel Xeon 64 bit CPU @3.60
RAM/GB	64
GPU	NVIDIA Quadro P4000, 8 GB, 1792 Cuda cores
GPU acceleration library	CUDA10.0, CUDNN7.4
TensorFlow	2.x
Keras	2.2.x

such as blur, flip, crop, affine, and contrast were applied to enhance the robustness of the model. The detection model was trained and tested using a workstation with the configurations listed in Table 1.

Model training was stopped at 100 epochs, as it attained the minimum training and testing loss accuracy of 0.12 and 0.45; at the same time, it reached training and testing accuracy of 98.3% and 96.8%, respectively, as shown in Figs. 4 and 5. The model was set to save the checkpoints (i.e., weight file of detection model on various parameters) for every 10 epochs. The size of the final weight file achieved is 12.4 MB, is a lightweight player detection network which can be embedded on low-edge devices for real-time implementation and may achieve better accuracy in detecting the player in real time.

Though the basketball match was captured using a dynamic camera, players were detected accurately, with the performance and robustness of the proposed method measured using four metrics whereas results were compared and tabulated as shown in Table 2. Figure 6 depicts that almost all the players were detected even though the movement of the camera changed while capturing the match, as shown between frame-8 to frame-199.

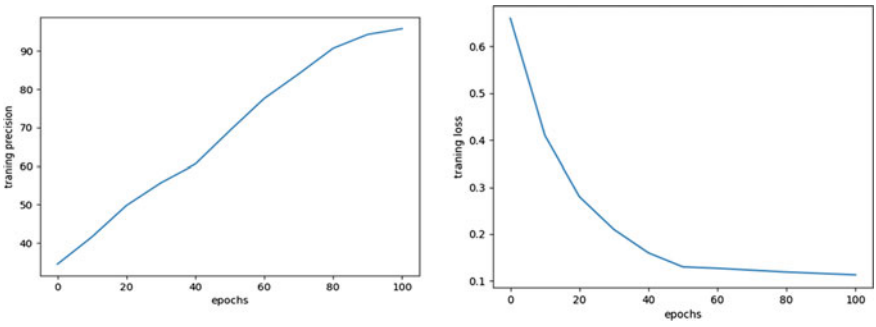


Fig. 4 Training precision and training loss with respect to number of epochs