Firat Univ Jour. of Exp. and Comp. Eng., 2(2), 46-54, 2023

Firat University Journal of Experimental and Computational Engineering
Fırat Üniversitesi Deneysel ve Hesaplamalı Mühendislik Dergisi

SKLBP14: A new textural environmental sound classification model based on a square-kernelled local binary pattern

SKLBP14: Kare çekirdekli yerel ikili modele dayalı yeni bir dokusal çevresel ses sınıflandırma modeli

Arif Metehan YILDIZ[1*], Mehmet Veysel GUN[2], Kubra YILDIRIM[3], Tugce KELES[4], Sengul DOGAN[5], Turker TUNCER[6], U. Rajendra ACHARYA[7,8,9]

[1,2,3,4,5,6]Department of Digital Forensics Engineering, College of Technology, Firat University, Elazig, Turkey.
[7]School of Mathematics, Physics and Computing, University of Southern Queensland, Springfield, Australia.
[8]Department of Biomedical Engineering, School of Science and Technology, SUSS University, Singapore.
[9]Department of Biomedical Informatics and Medical Engineering, Asia University, Taichung, Taiwan.
[1]a.metehanyildiz@gmail.com, [2]m.veyselgun@gmail.com, [3]kubra.yildirim@firat.edu.tr, [4]tkeles@firat.edu.tr, [5]sdogan@firat.edu.tr, [6]turkertuncer@firat.edu.tr, [7,8,9]Rajendra.Acharya@usq.edu.au

Received: 13.02.2023
Accepted: 22.03.2023
Revision: 25.02.2023
doi: 10.5505/fujece.2023.03521
Research Article

Abstract

Nowadays, the forward-forward (FF) algorithm is very popular in the machine learning society, and it uses a square-based activation function. In this research, we inspired the FF algorithm and presented a new kernel for a local binary pattern named square-kernelled local binary pattern (SKLBP). By deploying the proposed one-dimensional SKLBP, a new feature engineering model has been presented. To measure the classification ability of the proposed SKLBP-based model, we have collected a new textural environmental sound classification (ESC) dataset. The collected dataset is a balanced dataset, and it contains 15 classes. There are 100 sounds in each class. Our proposed model has mimicked the deep learning structure. Therefore, it uses multileveled feature extraction methodology by using discrete wavelet transform. The features generated have been considered as input for the iterative feature selector. The chosen feature vector has been utilized as input of the k nearest neighbor classifier. The proposed SKLBP-based signal classification model reached 94% classification accuracy. In this aspect, we contributed to the ESC methodology by collecting the new textural ESC dataset and proposing the SKLBP-based ESC model.

Keywords: Textural ESC dataset, Square-kernelled local binary pattern, Signal classification, Advanced signal processing, Textural feature extraction

Özet

Günümüzde ileri-ileri (FF) algoritması, makine öğrenimi toplumunda çok popülerdir ve kare tabanlı bir aktivasyon işlevi kullanır. Bu araştırmada, FF algoritmasından ilham aldık ve yerel ikili örüntü için yeni bir çekirdek sunduk ve bu, kare çekirdekli yerel ikili örüntü (SKLBP) olarak adlandırıldı. Önerilen tek boyutlu SKLBP'yi konuşlandırarak, yeni bir özellik mühendisliği modeli sunulmuştur. Önerilen SKLBP tabanlı modelin sınıflandırma yeteneğini ölçmek için, yeni bir dokusal çevresel ses sınıflandırması (ESC) veri seti topladık. Toplanan veri seti dengeli bir veri seti olup 15 sınıf içermektedir. Her sınıfta 100 ses vardır. Önerdiğimiz model derin öğrenme yapısını taklit etmiştir. Bu nedenle, ayrık dalgacık dönüşümü kullanarak çok düzeyli öznitelik çıkarma metodolojisini kullanır. Oluşturulan özellikler, yinelemeli özellik seçicinin girdisi olarak kabul edilmiştir. Seçilen öznitelik vektörü k en yakın komşu sınıflandırıcının girdisi olarak kullanılmıştır. Önerilen SKLBP tabanlı sinyal sınıflandırma modeli, %90'ın üzerinde doğruluğa ulaştı. Bu bağlamda, yeni dokusal ESC veri setini toplayarak ve SKLBP tabanlı ESC modelini önererek ESC metodolojisine katkıda bulunduk.

Anahtar kelimeler: Dokusal ESC veri seti, Kare çekirdekli yerel ikili model, Sinyal sınıflandırması, Gelişmiş sinyal işleme, Dokusal özellik çıkarma

*Corresponding author

1. Introduction

Today, machine learning techniques are used as decision support systems to assist experts in many different disciplines [1, 2]. These systems' main purpose is to reduce experts' workload or assist in areas where experts are not available [3]. For example, environmental sound classification (ESC) is a very popular topic, and different machine-learning techniques have been presented to classify environmental sounds [4, 5]. In this study, a new automatic ESC method is proposed. The dataset collected for this purpose has 15 classes. The proposed method and the results obtained are presented below.

1.1. Motivation and our proposal

Many researchers have proposed variable automatic ESC models [6-8]. Our primary motivations are:

- To propose a new textural feature extraction kernel and investigate the classification ability of this kernel,
- To present a new textural ESC dataset.

However, there are limited ESC models for textural sounds. To fill this literature gap, we gathered a new textural ESC dataset. This dataset has 1500 sounds with 15 classes. Moreover, a textural features-based feature engineering model has been presented in this study. The commonly known feature extraction function is a local binary pattern, and the advantages of the LBP [9, 10] are: (i) it is a very simple feature extractor, and any user can apply this feature extractor by deploying any programming environment, (ii) it generates salient/distinct features, (iii) 256 features are extracted without depending the length of the signal. However, LBP was proposed for images and images containing positive values. We have recommended a novel ESC model, and sounds contain both positive and negative values. Therefore, squared kernel has been introducred to use negative values as an actor for the feature creation by inspiring the FF algorithm, and our proposed feature extractor is named square-kernelled local binary pattern (SKLBP). SKLBP is a hand-crafted feature extractor, and hand-crafted feature extractors cannot create features at a high level. A multileveled feature creation/extraction architecture has been used to solve this problem. Herein, we have used 14 leveled structures. Thus, our methodology is named SKLBP14. SKLBP14 uses an iterative (loop-based) feature selector, and this feature selection function is iterative neighborhood component analysis (INCA). INCA can select the best feature combination among the used features. We used the k nearest neighbors (kNN) classifier in the classification phase with 10-fold cross-validation.

1.2. Innovations and Contributions

The contributions and innovations of this research are:
Innovations:

- A novel one-dimensional feature creation function has been recommended.
- A new textural environmental sound dataset was collected.
- A new generation feature engineering model has been proposed, and this model is named SKLBP14.

Contributions:

- Hinton [11] proposed a new machine-learning model, and it is named FF. FF uses a square-based activation function, and there is no need to use any backpropagation algorithm. The feature engineering models are generally forward models. Therefore, the FF algorithm can be beneficial for feature engineering. In this work, we only used the activation function of the FF algorithm as a kernel to show the benefits of the FF for feature engineering. In this aspect, we are the first team to use a square kernel for textural feature extraction.
- In the literature, there are variable ESC models and datasets. However, ESC is wide and specific ESC models should be presented. To contribute to this area, we acquired a new ESC dataset, a textural ESC dataset. This dataset is not a small dataset since there are 15 textural sounds in this dataset.
- Deep learning (DL) models are the flagships of ML. However, DL models are very expensive in view of the computational complexity [12, 13]. Feature engineering models must be used to propose lightweight models, but we should benefit from deep learning structures [14]. Hence, a new multilevel feature engineering model has been presented, and it is named SKLBP14.

of the foundation. For each distance, three different excavation was used with dimensions (1B, 1.5B and 2B).

2. Our Dataset

We collected a new textural sound dataset, a balanced dataset. There are 15 classes in this dataset, and the used classes are (i) applause, (ii) bubbling water, (iii) fire, (iv) frogs, (v) geese, (vi) helicopter, (vii) insects, (viii) jackhammer, (ix) rain, (x) rusting paper, (xi) sparrows, (xii) stream, (xii) train, (xiv) waves and (xv) wind. This dataset is a balanced dataset, and each class contains 100 sounds. The duration of each sound signal is about 2 seconds, and the sampling frequency of these sounds is 44.1 KHz. This textural sound dataset was collected from https://freesound.org/ and https://www.youtube.com/. This dataset contains textural sounds. Therefore, we named it TESC. This dataset is available at https://www.kaggle.com/datasets/senguldogan/texturalesc.

3. The Proposed SKLBP

Herein, we have presented the proposed SKLBP feature extractor. SKLBP feature extractor is a new version of the LBP. LBP was proposed for image classification, and a one-dimensional version of this selector has generally been used for signal classification. In the proposed SKLBP, we used the FF algorithm's activation function. This formula's main objective is to get the feature by using negative values. The schema of the SKLBP model is schematically explained in Figure 1.

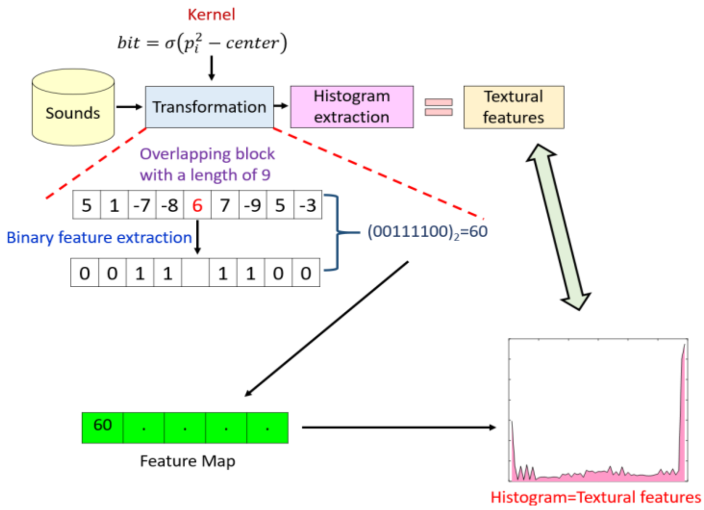

Figure 1. Graphical overview of the proposed SKLBP feature generator

In Figure 1, the schema of the proposed SKLBP feature extraction function has been given. This model is a LBP-like model and we have changed the kernel of this model to get distint fatures. The steps of these feature extractors are also listed below to clarify our proposed SKLBP feature extraction function. The steps are called S.

S1: Divide signal/sounds into overlapping blocks with nine lengths.

$$bl^i = S(i + j - 1), i \in \{1,2, \dots , l - 8\}, j \in \{1,2, \dots ,9\} \quad (1)$$

Herein, bl^i defines ith the created overlapping block with a length of nine, S defines the used signal, and l is the length of the used signal.

S2: Generate binary features from each overlapping block.

$$center = bl^i(5) \quad (2)$$

$$bit^i(k) = \sigma(bl^i(t)^2 - center), t \in \{1,2, \dots ,9\}\ and\ t \neq 5, k \in \{1,2, \dots ,8\} \quad (3)$$

$$\sigma(bl^i(t)^2 - center) = \begin{cases} 0, bl^i(t)^2 - center < 0 \\ 1, bl^i(t)^2 - center \geq 0 \end{cases} \quad (4)$$

The given equations above have been defined as our recommended binary feature generation method with our proposed square kernel. In these equations (Eqs 2-4), bit is a binary feature, σ is our used bit extraction function, and we used the square of the used non-center value in this function, and this function is similar to the signum function. In this step, 8 bits are created.

S3: Create a feature map signal with the calculated bits.

$$fm(i) = \sum_{k=1}^{8} bit^{i}(k) \times 2^{8-k} \quad (5)$$

where fm is the generated feature map signal.

S4: Extract histogram of the fm to generate feature vector with a length of 256 (=2^8).

$$feature = \phi(fm) \quad (6)$$

where $feature$ is the feature vector and ϕ is histogram extraction function.
The given four steps above (S1-S4) have been defined the proposed SKLBP feature extraction function.

4. The Proposed Model

In this section, the details of the proposed SKLBP model has been given. In this model, we mimicked DL structure. General phases of the proposed SKLBP14 are (i) SKLBP-based multileveled feature extraction, (ii) loop-based distinct feature selection and (iii) classification. To better explain our proposed SKLBP14 model, we schematized this model in Figure 2.

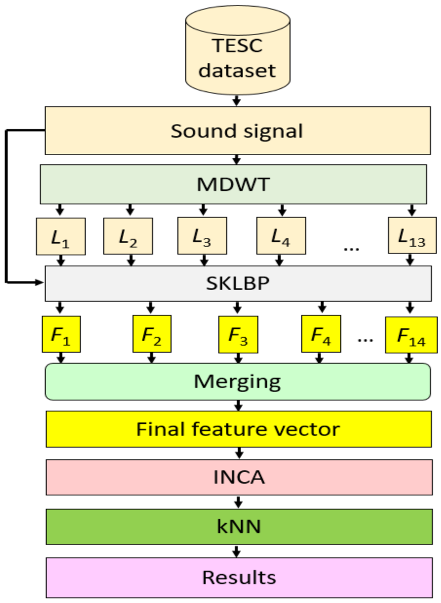

Figure 2. Schematic depiction of the proposed model. The abbreviations used in this figure are explained as follows

MDWT: multilevel discrete wavelet transform, L: low-pass filter wavelet coefficients (wavelet bands), F: feature vectors with a length of 256.

As stated in Figure 2, we have applied MDWT with 13 levels and used low-pass filter (L) bands since L bands are named approximation bands. By deploying SKLBP to these wavelet bands and the raw sound. We have generated features in both frequency and space domains by deploying this strategy. Moreover, a new multileveled feature generation method was by deploying MDWT. Herein, the mother wavelet function is symlet 4. Symlet 4 is a useful wavelet filter for sound noise reduction. Therefore, we selected this filter. By deploying the presented SKLBP to 14 signals (=13 wavelet bands + 1 sound), SKLBP14 has been created. In Figure 2, 14 feature vectors have been depicted using yellow box and F. The length of each feature vector is equal to 256. Therefore, there are 14 feature vectors. Hence, we have obtained 3584 (=256×14) features finally. The best feature combination among the generated 3584 has been selected by employing the INCA feature selector. For this research, the best 449 features were selected among the created 3584 features by INCA. The kNN classifier has classified these features. The steps of our SKLBP19 are:

Step 1: Read each sound from the collected TESC dataset.

Step 2: Deploy MDWT to generate wavelet bands. Symlet 4 has utilized a mother wavelet filter.

$$[L^1, H^1] = \psi(S) \tag{7}$$

$$[L^r, H^r] = \psi(L^{r-1}), r \in \{2,3, \dots ,13\} \tag{8}$$

Herein, L defines low-pass filter bands, H represents high-pass filter bands and ψ is the DWT function. By deploying Eqs. 7-8, 13 leveled MDWT has been created.

Step 3: Create feature vectors deploying the generated L bands, raw sound and SKLBP feature generator.

$$F^1 = \xi(S) \tag{9}$$

$$F^{q+1} = \xi(L^q), q \in \{1,2, \dots ,13\} \tag{10}$$

where F is the feature vector by generating the proposed SKLBP feature generator, and it is defined using ξ function. In this step, 14 feature vectors have been created.

Step 4: Merge the created features to get the merged/final feature vector.

$$X\big(a + 256 \times (b-1)\big) = F^b(a), a \in \{1,2, \dots ,14\}, b \in \{1,2, \dots ,256\} \tag{11}$$

where X is the final feature vector.

Step 5: Deploy NCA to the final feature vector created and get the qualified feature indexes.

$$id = \mathcal{N}(X, y) \tag{12}$$

where id is the sorted/qualified indexes, $\mathcal{N}$ is the NCA feature selection function, and y is the actual outputs.

Step 6: Select the feature vectors iteratively per the initial and final values of the loop.

$$cf^{r-sv+1}(d,i) = X\big(d, id(i)\big), i \in \{1,2, \dots ,r\}, r \in \{sv, sv+1, \dots , fv\}, d \in \{1,2, \dots , NoS\} \tag{13}$$

Herein, sv is the start value of the loop, fv is the last value of the loop, cf defines chosen feature vector, and NoS is the number of sounds.

Step 6: Compute the misclassification rates of the selected features. In this step, kNN has been utilized as a loss value calculator.

$$\mathcal{L}(r - sv + 1) = \mathcal{K}(cf^{r-sv+1}, y) \tag{14}$$

Herein, $\mathcal{L}$ is the loss value, and $\mathcal{K}$ is the kNN classifier. By deploying $\mathcal{K}$ to the feature vector, misclassification rates have been computed.

Step 7: Find/calculate the index of the minimum loss/misclassification value.

$$ix = min(\mathcal{L}) \tag{15}$$

where ix is the index of the minimum misclassification value.

Step 8: Choose the optimal features by deploying the index calculated in the previous step and chosen feature vectors.

$$sfv = cf^{ix-sv+1} \tag{16}$$

where sfv is the selected feature vector.

Step 9: Calculate results by applying kNN to the selected features.

$$R = \mathcal{K}(sfv, y) \tag{17}$$

Herein, R is the result.

These 9 steps above are defined our recommended SKLPB14.

5. Experimental Results and Discussions

The recommended model was programmed on the MATLAB (2021a) environment. A personal computer was used for implementation, and we coded this model functionally. The proposed model is a parametric architecture, and the parameters are given below.

MDWT: 13 leveled MDWT with symlet 4 filter [15].

SKLBP: Square kernel, 9 sized overlapping blocks, histogram-based feature extraction.

Feature generation: We applied SKLBP to 14 signals (1 sound + 13 wavelet bands).

INCA: Loop range: [100,1000], 901 feature vectors have been created. kNN with 10-fold cross-validation (CV) is considered a loss value calculator [16].

kNN: k is 1, distance is L1-norm (Manhattan), and validation is 10-fold CV [17, 18].

We used these parameters to implement the proposed model. In order to evaluate the performance of the presented SKLBP-based model, we have used accuracy, unweighted average precision (UAP), and overall F1 measures. We have used overall performance evaluation metrics since our dataset contains 15 classes. Moreover, we have presented class-by-class classification measures to show our SKLBP-based model's classification performance clearly. The evaluation metrics are mathematically defined below [19, 20].

$$Accuracy = \frac{tn + tp}{tn + fn + fp + tp} \tag{18}$$

$$Precision = \frac{tp}{tp + fp} \tag{19}$$

$$F1 = \frac{2tp}{2tp + fp + fn} \tag{20}$$

Herein, tn, fn, fp and tp are the number of true negatives, false negatives, false positives, and true positives, respectively. The calculated confusion matrix is demonstrated in Figure 3.

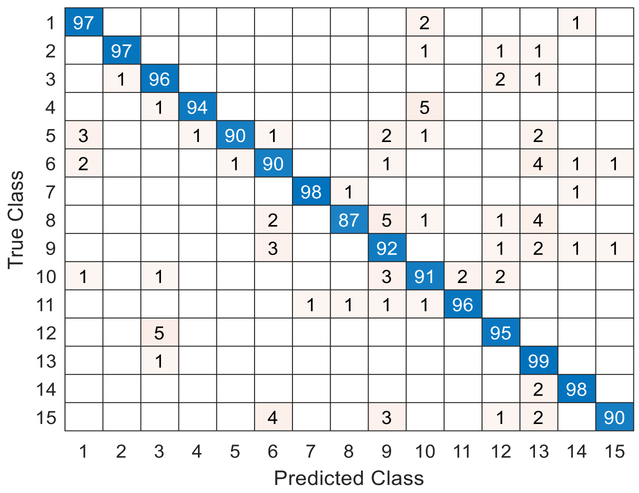

Figure 3. The calculated confusion matrix of the proposed SKLBP-based model

By using Figure 3, we have calculated overall performance metrics and these metrics are listed in Table 1.

Table 1. Results (%) of our recommended SKLBP-based textural ESC approximation

Performance metric	Result
Accuracy	94
UAP	94.26
F1	94.03

Table 1 tabulated that our proposal achieved over or equal 94% for all overall performance metric. Furthermore, we have calculated class-wise (class by class) classification performances by deploying confusion matrix and these results are depicted in Figure 4.

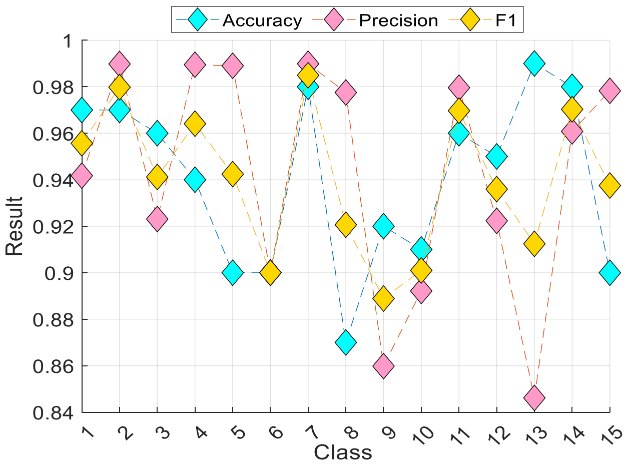

Figure 4. Class-wise classification performances

Figure 4 demonstrated that the best accurate class is the 13th class (it reached 99% class-wise accuracy), and this class contains the sound of the train, and the worst class is the 8th class (87% recall/class-wise accuracy was yielded), and this class is a jackhammer. Moreover, all class-wise results are higher than 84%.

The findings of this work are also listed below.

- We gathered a new textural ESC dataset. ESC is a commonly studied work but there are limited public textural ESC dataset. Therefore, we published this dataset to contribute the ESC.
- Forward-forward is a new and popular model in the machine learning society. A basic application (similar method) of the FF (the proposed SKLBP) for feature extraction has been presented.
- We have proposed a multilevel feature engineering model which is called SKLBP14 (we mimicked deep learning structures to get both low and high leveled features).
- Our proposed model attained 94% classification accuracy, 94.26% UAP and 94.03% overall F1-score. For this point, this model got high classification performance.
- 13th class is the best accurate class among the used 15 class since it attained 99% class-wise accuracy (recall).
- 10-fold cross-validation was used to get robust classification results.
- This model is a very simple model. Thus, other signal classification problems can easily solved by applying our proposed model.
- We have used hand-crafted feature extraction model and the time burden of this function is equal to $O(n)$. In this aspect, a lightweight ESC model has been presented.
- Soon, we are planning to gather more ESC sounds and we will propose a 1D-LBP-based deep learning model like convolutional neural network (CNN).

6. Conclusions

In this research, we have proposed a new kernel for LBP. This kernel is square-based, and this kernel is named SKLBP. By using SKLBP, we have presented a new feature engineering model. In order to test the classification performance our proposal, we have collected a new TESC dataset. This dataset contains 1500 sounds with 15 classes. In this model, MDWT was applied to generate frequency bands, and SKLBP extracted features from bands and the raw sound. The generated features were fed to the INCA feature selector, and the most informative features were selected. Finally, the selected features were utilized as input for the kNN classifier to get classification results. By deploying our model, 94% classification accuracy was attained. This result clearly highlighted that our presented SKLBP-based sound classification model is good for textural sound classification.

Collecting a bigger TESC dataset and presenting an explainable feature engineering model for TESC are among our near future plans. Moreover, the presented SKLBP-based model can be applied to other one-dimensional data classification problems.

7. Author Contribution Statement

The authors have no conflicts of interest to declare. All co-authors have seen and agree with the contents of the manuscript and there is no financial interest to report. We certify that the submission is original work and is not under review at any other publication.

8. Ethics Committee Approval and Conflict of Interest

"There is no conflict of interest with any person/institution in the prepared article"

9. References

[1] Demir K, Berna A, Demir F. "Detection of brain tumor with a pre-trained deep learning model based on feature selection using MR images". *Firat University Journal of Experimental and Computational Engineering,* 2(1), 23-31, 2023.

[2] Sevinç A, Özyurt F. "Classification of recyclable waste using deep learning architectures". *Firat University Journal of Experimental and Computational Engineering,* 1(3), 122-128, 2022.
[3] Matwin S, Kouznetsov A, Inkpen D, Frunza O, O'Blenis P. "A new algorithm for reducing the workload of experts in performing systematic reviews". *Journal of the American Medical Informatics Association,* 17(4), 446-453, 2010.
[4] Salamon J, Bello JP. "Deep convolutional neural networks and data augmentation for environmental sound classification". *IEEE Signal processing letters,* 24(3), 279-283, 2017.
[5] Abdoli S, Cardinal P, Koerich AL. "End-to-end environmental sound classification using a 1D convolutional neural network". *Expert Systems with Applications,* 136, 252-263, 2019.
[6] Okaba M, Tuncer T. "An automated location detection method in multi-storey buildings using environmental sound classification based on a new center symmetric nonlinear pattern: CS-LBlock-Pat". *Automation in Construction,* 125, 103645, 2021.
[7] Zhang Y, Zeng J, Li Y, Chen D. "Convolutional neural network-gated recurrent unit neural network with feature fusion for environmental sound classification". *Automatic Control and Computer Sciences,* 55, 311-318, 2021.
[8] Demir F, Abdullah DA, Sengur A. "A new deep CNN model for environmental sound classification". *IEEE Access,* 8, 66529-66537, 2020.
[9] Ojala T, Pietikainen M, Maenpaa T. "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns". *IEEE Transactions on pattern analysis and machine intelligence,* 24(7), 971-987, 2002.
[10] Ahonen T, Hadid A, Pietikäinen M. "Face recognition with local binary patterns". in *European conference on computer vision*, 2004: Springer, 469-481.
[11] Hinton G. "The forward-forward algorithm: Some preliminary investigations". *arXiv preprint arXiv:2212.13345*, 2022.
[12] Kotsiopoulos T, Sarigiannidis P, Ioannidis D, Tzovaras D. "Machine learning and deep learning in smart manufacturing: The smart grid paradigm". *Computer Science Review,* 40, 100341, 2021.
[13] Hall O, Dompae F, Wahab I, Dzanku FM. "A review of machine learning and satellite imagery for poverty prediction: Implications for development research and applications". *Journal of International Development,* 2023.
[14] Tuncer T, Dogan S, Baygin M, Acharya UR. "Tetromino pattern based accurate EEG emotion classification model". *Artificial Intelligence in Medicine,* 123, 102210, 2022.
[15] Tuncer T, Dogan S, Pławiak P, Acharya UR, "Automated arrhythmia detection using novel hexadecimal local pattern and multilevel wavelet transform with ECG signals". *Knowledge-Based Systems,* 186, 104923, 2019.
[16] Tuncer T, Dogan S, Özyurt F, Belhaouari SB, Bensmail H. "Novel Multi Center and Threshold Ternary Pattern Based Method for Disease Detection Method Using Voice". *IEEE Access,* 8, 84532-84540, 2020.
[17] L. E. Peterson, "K-nearest neighbor," *Scholarpedia,* vol. 4, no. 2, p. 1883, 2009.
[18] Maillo J, Ramírez S, Triguero I, Herrera F. "kNN-IS: An Iterative Spark-based design of the k-Nearest Neighbors classifier for big data". *Knowledge-Based Systems,* 117, 3-15, 2017.
[19] Carrington AM. *et al.* "Deep ROC analysis and AUC as balanced average accuracy, for improved classifier selection, audit and explanation". *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 45(1), 329-341, 2022.
[20] Powers DM. "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation". *arXiv preprint arXiv:2010.16061*, 2020.