ORIGINAL ARTICLE

Swin-LBP: a competitive feature engineering model for urine sediment classification

Mehmet Erten[1] · Prabal Datta Barua[2] · Ilknur Tuncer[3] · Sengul Dogan[4] · Mehmet Baygin[5] · Turker Tuncer[4] · Ru-San Tan[6,7] · U. Rajendra Acharya[8]

Abstract
Automated urine sediment analysis has become an essential part of diagnosing, monitoring, and treating various diseases that affect the urinary tract and kidneys. However, manual analysis of urine sediment is time-consuming and prone to human bias, and hence there is a need for an automated urine sediment analysis systems using machine learning algorithms. In this work, we propose Swin-LBP, a handcrafted urine sediment classification model using the Swin transformer architecture and local binary pattern (LBP) technique to achieve high classification performance. The Swin-LBP model comprises five phases: preprocessing of input images using shifted windows-based patch division, six-layered LBP-based feature extraction, neighborhood component analysis-based feature selection, support vector machine-based calculation of six predicted vectors, and mode function-based majority voting of the six predicted vectors to generate four additional voted vectors. Our newly reconstructed urine sediment image dataset, consisting of 7 distinct classes, was utilized for training and testing our model. Our proposed model has several advantages over existing automated urinalysis systems. Firstly, we used a feature engineering model that enables high classification performance with linear complexity. This means that it can provide accurate results quickly and efficiently, making it an attractive alternative to time-consuming and biased manual urine sediment analysis. Additionally, our model outperformed existing deep learning models developed on the same source urine sediment image dataset, indicating its superiority in urine sediment classification. Our model achieved 92.60% accuracy for 7-class urine sediment classification, with an average precision of 92.05%. These results demonstrate that the proposed Swin-LBP model can provide a reliable and efficient solution for the diagnosis, surveillance, and therapeutic monitoring of various diseases affecting the kidneys and urinary tract. The proposed model's accuracy, speed, and efficiency make it an attractive option for clinical laboratories and healthcare facilities. In conclusion, the Swin-LBP model has the potential to revolutionize urine sediment analysis and improve patient outcomes in the diagnosis and treatment of urinary tract and kidney diseases.

Keywords Urinary analysis · Urinary sediment classification · Swin-LBP · Biomedical image classification

1 Introduction

Microscopic urine sediment analysis is a routine laboratory test [1]. Identifiable urine sediments include erythrocytes, leukocytes, crystals, casts, epithelial cells, sperms, bacteria, and mycetes (fungi) [2, 3], which, either singly or in combination, can connote the presence of different clinical conditions [4]. For example, the presence of erythrocyte urinary sediments above a specified threshold means that there is bleeding into the urine, which may signify diverse

pathologies (stone, infection, cancer, etc.) affecting anatomical structures along the urinary tract, e.g., kidney glomerulus, kidney tubules, ureter, bladder, prostate, and urethra. Here, the shape of urinary erythrocytes depends on the origin, i.e., dysmorphic versus non-dysmorphic in glomerular and non-glomerular hematuria, respectively [2]. Urine sediment microscopic examination can be conducted manually, labor-intensive and subject to human bias [5] or using automated devices. The latter enhances operation efficiency, helps reduce laboratory work burden, and is an invaluable diagnostic screening tool in high-volume clinical laboratories [6]. Indeed, image-based intelligent

Extended author information available on the last page of the article

analytic systems can offer accurate and robust results at high throughput [7, 8] for diagnosis, surveillance, and therapeutic monitoring of various kidney and urinary tract diseases [9].

Automated image recognition is integral to urinalysis automation and generally comprises segmentation [10], feature selection [11], optimization [12, 13], and classification steps [14, 15]. Wide variations in sediment shapes, small cell sizes, and occasional clumping of cells pose challenges to the application of machine learning for urinary sediment classification [7, 14]. Table 1 summarizes the state-of-the-art image-based urine sediment analysis, which comprises deep learning models exclusively. Of note, large numbers of images (ranging from a few to 300 thousand) of distinct sediment types (3–10 classes) have been studied that collectively encompassed the sediment types commonly encountered in the laboratory: red blood cells, white blood cells, epithelial cells, hyaline casts, mucus strands, crystals (e.g., calcium oxalate), spermatozoa, as well as exogenous infectious agents like bacteria and yeast cells.

It can be noted from Table 1 that the literature gaps of the urinary sediment classification models are listed below.

- To achieve high classification accuracies, deep learning models have typically been employed, despite their high computational complexity. However, there is a need for lightweight models that can run on simpler configurations such as a laptop.
- Datasets with a limited number of categories or with large distances between categories, such as RBC and sperm cells, may require alternative approaches to achieve high accuracy.
- All models listed in Table 1 are deep learning models, highlighting the need for the proposal of competitive feature engineering models as an alternative to deep learning models.

The main objective of this study is to address the gaps in feature engineering by proposing a new architecture, i.e., Swin-LBP in this work. This architecture represents a new generation feature engineering work that we believe will lead to improved performance in a variety of image analysis tasks. Furthermore, we have also developed a novel dataset of urinary cell images, which includes 7 different urinary cell categories. This dataset is an important contribution to the field, as it will allow researchers to benchmark the performance of their algorithms on a standardized dataset. Also, the created dataset contains more than 10,000 images.

1.1 Motivation and our method

We were motivated to develop an accurate handcrafted urine sediment classification method based on computer vision of urine cell images. The challenge was posed as an image classification problem, to which deep learning networks have been widely applied [20–22]. Computer vision, which uses convolution neural networks (CNN) and transformer-based models [23, 24], has emerged as an important tool for image-based classification. Transformers-based models possess interesting architecture [25]. The popular vision transformer [26] and Swin transformer [27] rely on patch-based classification: the former uses fixed-sized patches to extract deep features, and the latter four-layered shifted windows patch division. In this work, we proposed a hand-modeled image classification method based on the Swin architecture in combination with a local binary pattern (LBP) [28] feature extraction function, which we named Swin-LBP. The model comprised five phases—(i) preprocessing of input images using shifted windows-based patch division; (ii) LBP-based feature extraction [28]; (iii) neighborhood component analysis (NCA) [29]-based feature selection; (iv) support vector machine (SVM) [30]-based classification; and (v) majority

Table 1 Recent image-based models developed for automated urinary sediment analysis

Author	Number of images/classes	Method
Liang [14]	5377; 7: RBC, WBC, EP, EP nuclei, cast, crystals, mycete	DenseNet features a pyramid network
Pan [16]	Unknown; 3: RBC, WBC, CAOX	Custom CNN
Li [9]	2551; 4: RBC, WBC, crystal, EP	Custom CNN based on LeNet-5 structure
Ji [17]	300,000; 10: RBC, WBC, WBC cluster, squamous EP, sperm, CAOX, hyaline cast, mucous filament, BACT, yeast	Data augmentation, area feature algorithm, AlexNet-based end-to-training CNN
Velasco [18]	1671; 3: RBC, CAOX, BACT	InceptionResNetV2; InceptionV3
Liu [19]	2024; 7: RBC, WBC, EP, cast, crystal, yeast, non-cell	Fine-tuned CNN ensemble: ResNet50, GoogleNet, De-AlexNet

BACT bacteria; *CAOX* calcium oxalate crystal; *CNN* convolutional neural network; *EP* epithelial cell; *RBC* red blood cell; *Rec* recall; *Sen* sensitivity; *Spe* specificity; *WBC* white blood cell

voting—and was trained and tested on a new reconstructed 7-class urine sediment image dataset.

1.2 Contributions

The contributions of the proposed model are given below.

- Novel handcrafted transformer-inspired Swin feature engineering model is proposed.
- We have built an effective and efficient computer vision model by combining:

  - Shifted windows-based patch division of images.
  - Computationally lightweight handcrafted feature extraction and selection functions.
  - Standard shallow classifier.
  - A simple majority voting algorithm has been used to get general classification results.

- Trained and tested on a 7-class urine sediment image dataset, the Swin-LBP model attained salutary 7-class classification accuracy of 92.60%.

The given bullets demonstrated that we are the first team to use swin architecture for handcrafted features. It is a new way to get high classification results from shallow methods.

2 Dataset

Utilizing a published dataset [14, 31, 32], we conducted segmentation, extraction, and cropping of individual urine sediment images. This process yielded a collection of 12,330 images, which we subsequently grouped into seven distinct classes: (i) cast (inclusive of all types of casts); (ii) crystal; (iii) epithelia; (iv) epithelial nuclei; (v) erythrocyte; (vi) leukocyte; and (vii) mycete. This dataset contains more than 10,000 images, and the distribution of the aforementioned dataset is tabulated in Table 2.

We have randomly selected 2000 images from each category but there are lower than 2000 observations in the crystal and epithelial nuclei datasets. Therefore, all images from these datasets have been involved in our used urine cell image dataset. Sample images about this dataset are also demonstrated in Fig. 1.

3 Swin-LBP model

Our novel contribution to the field of computer vision is a cutting-edge model that we call Swin-LBP. The main objective of this work is to significantly improve the classification capabilities of shallow models. As depicted in

Table 2 Attributes of the collected dataset

No	Name	Number of images
1	Cast	2000
2	Crystal	1643
3	Epithelia	2000
4	Epithelial nuclei	687
5	Erythrocyte	2000
6	Leukocyte	2000
7	Mycete	2000
Total		12,330

Fig. 2, our proposed model comprises five phases that are designed to work together seamlessly.

The first phase involves preprocessing the data, whereby each urine sediment image is resized to 240×240 and then it is divided into *six* patches of various sizes (30×30, 40×40, 48×48, 60×60, 80×80, and 120×120). This process generates 64, 36, 25, 16, 9, and 4 patches for each resized image. Next, in the second phase, we use LBP [28] to extract 59 features (as described in Sect. 3.2) from each of the patches and the undivided sediment images. This results in a large number of feature vectors that are one more than the number of patches at every extraction layer. To handle this, we merge the generated feature vectors to create six merged feature vectors for each input sediment image.

In the third phase, we employ the neighborhood component analysis (NCA) [29] feature selection function to select the most informative 295 features from each feature vector, thereby balancing the lengths of the six feature vectors. In the fourth phase, we feed the six selected vectors, each containing the top discriminative features, to a shallow support vector machine (SVM) [30] classifier to obtain six predicted vectors using tenfold cross-validation strategy.

Finally, in the fifth and last phase, we apply a majority voting algorithm to the six predicted vectors to obtain four predicted vectors. From the six predicted vectors and four voted vectors obtained in the fourth and fifth phases, respectively, we select the one with the most accurate result as the final output. We provide technical details of each phase in the following sections. Moreover, we have illustrated a block diagram (open version) of the proposed Swin-LBP in Fig. 3.

3.1 Preprocessing

In this first phase, Swin architecture-inspired shifted windows-based patch division is performed as follows:

Fig. 1 Sample urine cell images of the used dataset **a** cast, **b** crystal, **c** epithelia, **d** epithelia nuclei, **e** erythrocyte, **f** leukocyte, **g** mycete
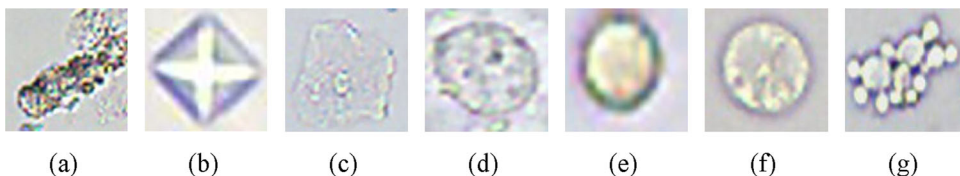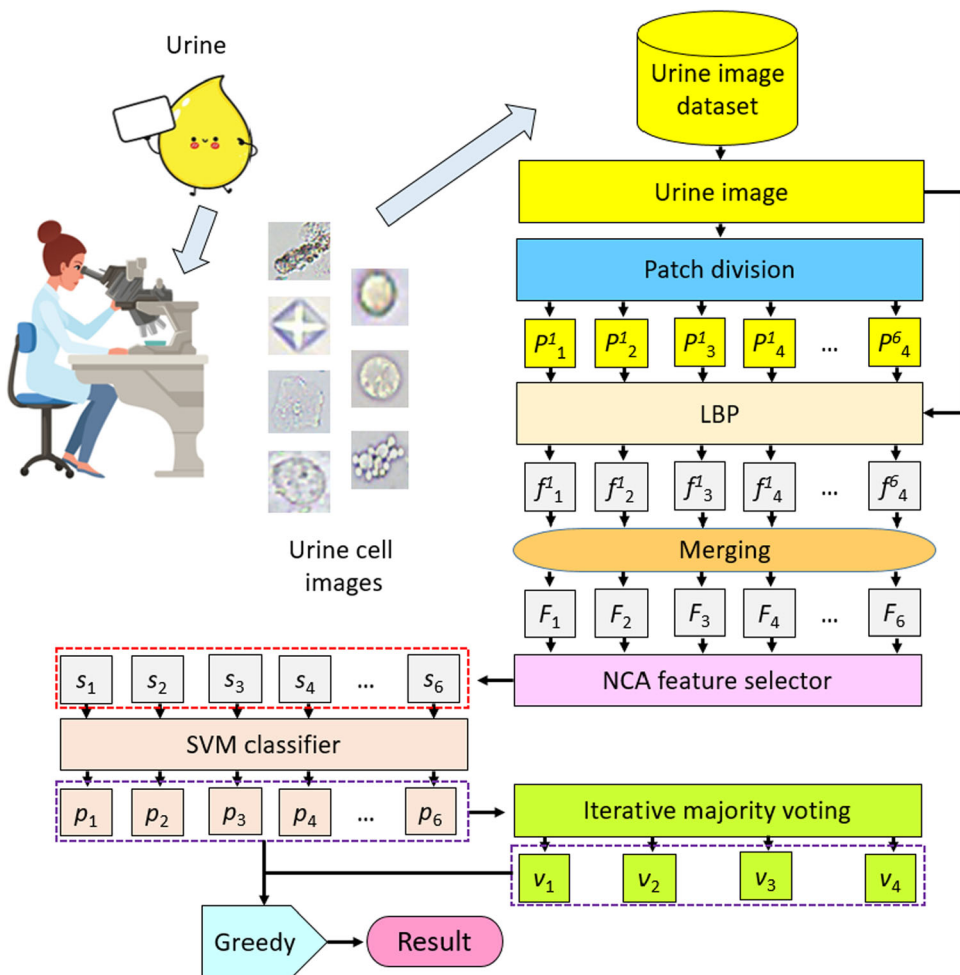

(a)   (b)   (c)   (d)   (e)   (f)   (g)

Fig. 2 Graphical depiction of the proposed model. P: patches; LBP: local binary pattern; f: the generated individual feature vectors; F: merged feature vectors; s: selected feature vectors; p: predicted labels; v: voted vectors

Step 0: Read urine sediment images from the collected dataset.

Step 1: Resize each image to a 240 × 240 sized image.

Step 2: Apply six types of patch division to create six layers. This process is defined below.

$$p_t^k = Im(i : i + s_n - 1, j : j + s_n - 1),$$
$$s \in \{30, 40, 48, 60, 80, 120\}$$
$$k \in \{1, 2, \ldots, 6\}, t \in \left\{1, 2, \ldots, \frac{240}{s_n}\right\}, \quad (1)$$
$$i \in \{1, s_n, \ldots, 240\}, j \in \{1, s_n, \ldots, 240\}$$

where p represents patch; Im, the used image; k, the type of patch; s, the size of the patch; and t, the number of patches.

3.2 Feature extraction

LBP is a histogram-based feature extraction function deployed in the model to extract global and local textural features from the undivided resized sediment image and its corresponding patches, respectively, using neighborhood relations constrained within microstructural image units of 3 × 3 overlapping windows (Fig. 4).

For each resized input sediment image, the extract LBP function of MATLAB is used to extract at every one of the six layers of patch divisions 59 features from the undivided image, and its corresponding derived patches.

Step 3: Extract features from the resized images and generate patches. This process is defined below.

Fig. 3 Block diagram of the Swin-LBP model (see text for detailed description). f, extracted feature vector; P, patch, SVM, support vector machine; LBP, local binary pattern; NCA, neighborhood component analysis
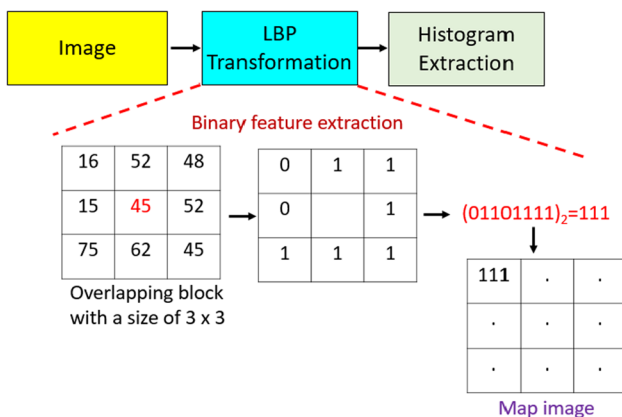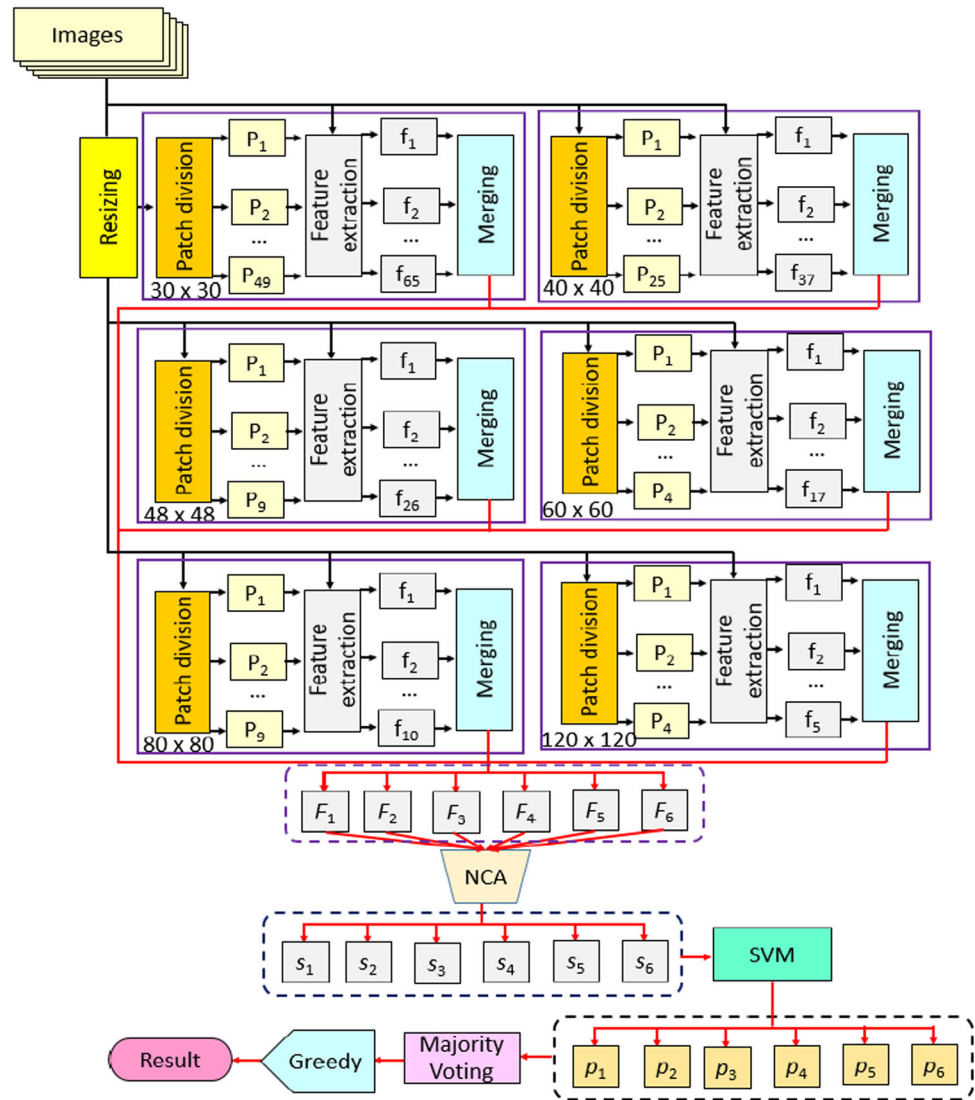

Fig. 4 Block diagram of the LBP feature extraction function

$$f_1^k = bp(Im)$$
$$f_{t+1}^k = bp(p_t^k)$$

(2)

where f represents the generated feature vector with a length of 59; and $bp(.)$, LBP function.

Step 4: Merge the generated feature vectors in every layer to create six merged feature vectors per input image.

$$F^k(j + 59 \times (h-1)) = f_t^k(j), j \in \{1, 2, \ldots, 59\}, h \in \{1, 2, \ldots, t+1\}$$

(3)

where F^k represents the kth merged feature vector; the lengths of F^1, F^2, F^3, F^4, F^5, and F^6 being 3853 (= 65 × 59), 2183 (= 37 × 59), 1534 (= 26 × 59), 1003 (= 17 × 59), 590 (= 10 × 59), and 295 (= 5 × 59), respectively.

3.3 Feature selection

The NCA function, a simple and effective L1-norm distance-based feature selector [29], is deployed to select the most discriminative 295 features in each of the merged six feature vectors, which are unequal in lengths (see Sect. 3.2), generated per input urine sediment image. In so doing, two important aims are achieved: (i) reduction in data dimensionality; and (ii) balancing/equalizing the lengths of the resultant NCA-selected feature vectors to 295.

Step 5: Apply qualified indexes using the NCA feature selection function.

$$ind^k = \xi(F^k, y) \tag{4}$$

where $\xi(.,.)$ represents the NCA feature selection function; ind^k implies the qualified indexes of the features; and y defines actual labels.

Step 6: Choose the most informative 295 features from the extracted feature vectors.

$$s^k(d, i) = F^k\big(d, ind^k(i)\big), d \in \{1, 2, \ldots, NoI\}, i \in \{1, 2, \ldots, 295\} \tag{5}$$

where s^k represents the kth feature vector with a length of 295 and NoI, the number of images.

3.4 Classification

A shallow classifier has been used in this model, and this classifier is SVM. SVM is the widely used classifier in the literature. Hyperparameters are set at: kernel function is polynomial, polynomial order is three, box constraint is 1, coding one-vs-all, validation is tenfold cross-validation. The classification process is defined below.

Step 7: Apply SVM-based classification.

$$p^k = \kappa(s^k, y) \tag{6}$$

where p represents the predicted vector; and $\kappa()$, the SVM classifier function.

3.5 Majority voting

Mode function-based weightless/hard majority voting is implemented to augment the classification performance of the Swin-LBP model. This process is defined below.

Step 8: Calculate the accuracies of the predicted vectors.

Step 9: Sort predicted vectors in descending order of accuracy rates.

$$id = sort(p^k) \tag{7}$$

where id represents the index of the sorted vector.

Step 10: Generate four voted predicted vectors.

$$v^1 = \omega\left(p^{id(1)}, p^{id(2)}, p^{id(3)}\right) \tag{8}$$

$$v^2 = \omega\left(p^{id(1)}, p^{id(2)}, p^{id(3)}, p^{id(4)}\right) \tag{9}$$

$$v^3 = \omega\left(p^{id(1)}, p^{id(2)}, p^{id(3)}, p^{id(4)}, p^{id(5)}\right) \tag{10}$$

$$v^4 = \omega\left(p^{id(1)}, p^{id(2)}, p^{id(3)}, p^{id(4)}, p^{id(5)}, p^{id(6)}\right) \tag{11}$$

where v represents the voted vector, and $\omega()$, the mode function.

Step 11: Calculate accuracies of the four voted predicted vectors and select the most accurate voted one.

4 Results

Two performance evaluation metrics were used to evaluate the model: accuracy and F1-score [33, 34]. These equations are given below.

$$acc = \frac{tp + tn}{tp + tn + fp + fn} \tag{12}$$

$$f1 = \frac{2tp}{2tp + fp + fn} \tag{13}$$

where acc represents accuracy; $f1$, F1-score; and tp, tn, fp, and fn, the number of true positives, true negatives, false positives, and false negatives, respectively. The performances of the proposed model have been presented using a tenfold cross-validation strategy.

4.1 Results of each layer

The Swin-LBP model extracts feature from six layers, each with variable defined patch sizes, input to downstream NCA feature selector and SVM classifier. Table 3 summarizes layer-wise classification performance obtained by the SVM classifier.

Table 3 7-class support vector machine-based classification performance by feature extraction layer

Layer	Patch size	Accuracy (%)	F1-score (%)
Layer 1	30 × 30	88.87	87.05
Layer 2	40 × 40	89.69	87.90
Layer 3	48 × 48	89.60	87.73
Layer 4	60 × 60	90.69	89.08
Layer 5	80 × 80	90.67	88.90
Layer 6	120 × 120	89.52	87.53

4.2 Voted results

The mode function used a majority voting algorithm to select the best performance among the layer-wise predicted vectors. Table 4 summarizes the voted results obtained by hard voting. The Swin-LBP model attained the highest classification accuracy of 92% with the third voted vector, which is higher than the best layer-wise accuracy of 90.69% (Layer 4) (Table 3) attained before majority voting. Our proposed Swin-LBP was designed as a self-organized image classification architecture. Therefore, the most accurate result among the ten (six layer-wise plus four voted results) was selected as the final result, i.e., the Swin-LBP model attained 92.60% classification accuracy (and 91.19% overall F1-score) on the urine sediment image dataset.

4.3 Class-wise results

Figure 5 and Table 5 depict the model's confusion matrix, and class-wise performance, respectively, based on the final best results, which were determined by the highest accuracy scores among the six layer-wise and four voted results calculated by the SVM classifier and majority voting, respectively. The best and worst classification performances were obtained for the "erythrocyte" (96.15% accuracy; 95.93% F1-score) and "epithelial nuclei" (69.26% accuracy; 76.96% F1-score) urine sediment classes, respectively.

4.4 Time complexity analysis

Swin-LBP is a lightweight feed-forward image classification model based on handcrafted feature extraction. Using big O notation, the model complexity is shown to be linear (Table 6).

5 Discussion

In this paper, a new urinalysis classification model was proposed that was trained and tested on a urine sediment image dataset comprising 6687 urine sentiment images
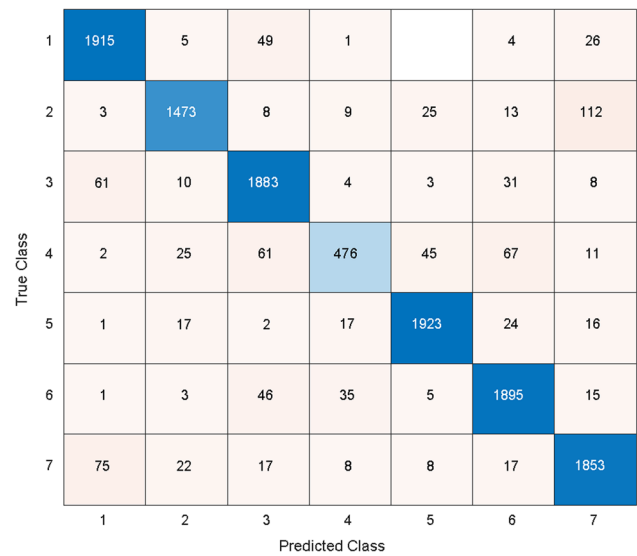

Fig. 5 Confusion matrix obtained by applying the model on the urine sediment image dataset. Classes 1 to 7 correspond to the urine sediment classes "Cast," "Crystal," "Epithelia," "Epithelial nuclei," "Erythrocyte," "Leukocyte," and "Mycete," respectively

Table 5 Model classification performance for each urine sediment class

Urine sediment class	Accuracy (%)	F1-score (%)
Cast	95.75	94.38
Crystal	89.65	92.12
Epithelia	94.15	92.62
Epithelial nuclei	69.26	76.96
Erythrocyte	96.15	95.93
Leukocyte	94.75	93.56
Mycete	92.65	91.71

divided equally among seven distinct classes. The novel Swin-LBP model employed a new learning architecture inspired by a Swin transformer in combination with a handcrafted LBP-based feature extractor. Despite its linear time complexity, our Swin-LBP model attained 92.60% classification accuracy for 7-class classification problem which is commensurate with the classification performance of more computationally demanding CNN-based deep learning models that had been developed on the same dataset from which our study dataset was derived (Table 7).

It can be noted from Table 7 that our proposed model reached high classification performance and it is a competitive feature engineering model to deep learning model.

Table 4 7-class model classification performance obtained by majority voting on the predicted feature vectors

Vector	The used layers	Accuracy (%)	F1-score (%)
Voted 1	4, 5, 2	92.07	90.63
Voted 2	4, 5, 2, 3	92.11	90.66
Voted 3	4, 5, 2, 3, 6	92.55	91.66
Voted 4	4, 5, 2, 3, 6, 1	92.60	91.19

Table 6 Computational complexity of the Swin-LBP model

Phase	Time complexity
Preprocessing (resizing; patch division)	$O(w \times h)$
Feature extraction and merging	$O(m \times n \times t + t)$
Feature selection	$O(k)$
Classification	$O(d)$
Majority voting	$O(l)$
Total	$O(w \times h + m \times n \times t + k + d + l)$

w and h represent width and height, respectively, of the input urine cell images; m and n, width and height, respectively, of the patches; t, the number of patches; k and d, time complexity coefficients of neighborhood component analysis feature selector and support vector machine classifier, respectively; and l, number of validation prediction vectors

Table 7 Comparison of automated image-based models for urine sediment analysis using the same dataset or its derivative*

Author	Method	Number of images/classes	Results (%)
Kang [35]	CNN-based image classification	5376 images; 7 classes	AP: 84.10
Liang [14]	DenseNet feature pyramid network	5377 images; 7 classes	AP: 86.90
Liang [31]	Faster region-based CNN; single shot detector	5376 images; 7 classes	AP: 84.10
Yan [32]	Bidirectional context propagation network	5377 images, 7 classes	AP: 88.20
Our model	Shifted windows patch division, local binary pattern, neighbor component analysis, support vector machine, majority voting	12,230 urine sediment images*; 7 classes	AP: 92.05 Acc: 92.60 F1: 91.19

Acc accuracy; *AP* average precision; *CNN* convolutional neural network; *F1* F1-score

*The images are reconstructed single sediment images extracted from the dataset used in the other four studies

5.1 Ablations

We have presented the ablations to show the high classification performance of the presented Swin-LBP model. In the first phase of this section, we used the shallow classifiers to get comparative results. Then, we compared results to LBP and local phase quantization (LPQ)-based models. These items have been defined below.

> *Item 1:* We have used LBP, histogram-oriented gradients (HOG), and local phase quantization (LPQ) feature extractors to get feature vectors.
> *Item 2:* neural network (NN), k-nearest neighbors (k-NN), and linear discriminant (LD) have been utilized as classifiers.
> *Item 3:* We have used a tenfold cross-validation strategy to get results. In this item, threefold cross-validation and fivefold cross-validation have been used and the calculated results using these validations have been presented for comparison.

The results of this ablation study are illustrated in Fig. 6.

According to Fig. 6, the proposed Swin-LBP achieved a classification accuracy of 92.60% using the utilized dataset. We used LBP as the primary feature extraction function for the introduced Swin-LBP architecture, resulting in an 80.84% classification accuracy for the LBP-based urinary image classification model. Therefore, our proposed Swin-LBP outperforms LBP by 11.76% points for this dataset. In addition, HOG and LPQ feature extraction-based models achieved 85.77% and 86% accuracies, respectively. These results suggest that Swin-LBP is the optimal handcrafted model among all considered models.

We used NN, k-NN, LD and SVM classifiers to get benchmark results. To attain classification accuracy, we have used feature vectors of layer 4 (60×60 sized patches) and the classification accuracies of these classifiers are shown in Fig. 7.

In Fig. 7, we observe that the SVM classifier achieves the highest accuracy of 90.63%, surpassing the NN classifier's performance of 87.20%.

In the defined Item 3, the results obtained using three cross-validations techniques (threefold CV, fivefold CV and tenfold CV) for the generated six feature vectors are shown in Fig. 8.

Figure 8 demonstrates that the tenfold cross-validation (CV) technique produces the highest accuracy among the validation methods used. However, we also evaluated the proposed Swin-LBP approach using two additional
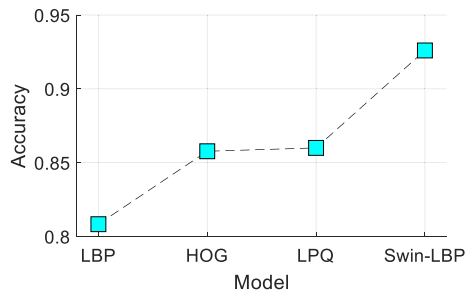
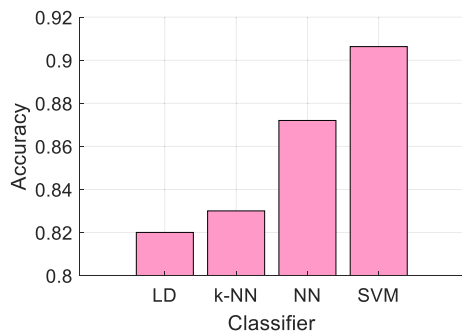**Fig. 6** Classification accuracies obtained for the ablation study

Fig. 7 Accuracy obtained using various classifiers

Fig. 8 Classification accuracies obtained for various feature vectors and SVM classifier with three validations (threefold, fivefold and tenfold CVs)

Fig. 9 Summary of accuracies obtained for three cross-validation techniques

validation techniques. Furthermore, we have calculated the final results using information fusion (iterative majority voting) of the used k-fold CV-based models and these classification results have been depicted in Fig. 9.

Figure 9 depicts that the tenfold CV is the best validation technique.

Our experimental results for ablations, as shown above, validate the effectiveness of the proposed Swin-LBP approach in achieving the optimal combination of parameters.

5.2 Highlights of the study

We have given highlights of this study in three following items: (1) findings, (2) advantages and (3) limitations. These important points have been listed below.

5.2.1 Findings

- Presents an automated urine sediment analysis system.
- Swin-LBP model uses machine learning algorithms to classify urine sediment images.
- It is based on Swin transformer architecture and LBP feature extraction technique and has five phases. Moreover, the introduced Swin-LBP has six feature extraction layers to use six types of fixed-size patch divisions.
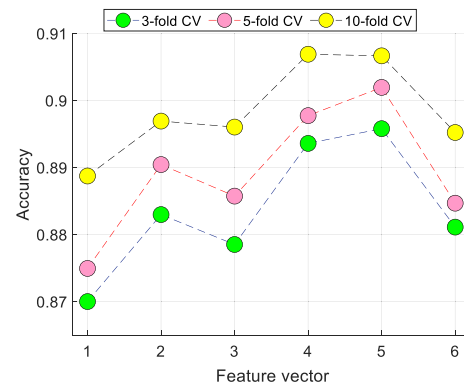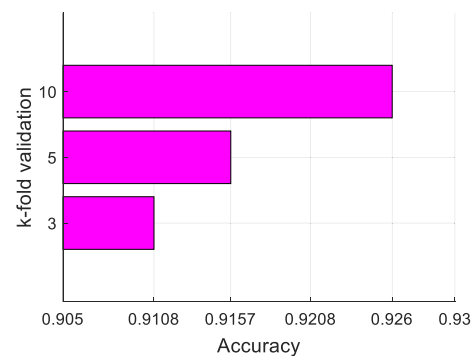
- The phases are preprocessing, feature extraction, feature selection, support vector machine-based calculation, and majority voting.
- The model was trained and tested on a 7-class urine sediment image dataset containing 12,330 images. This model achieved an accuracy of 92.60% and an average precision of 92.05%.
- According to class-wise results, the best accurate cell type is Erythrocyte and the worst one is epithelial nuclei. There are only 687 epithelial nuclei cell images in this dataset, and these cells are similar to epithelial urine cells.
- The best feature extraction layer is the 4th feature extraction layer. We have used 60×60 sized patches in this feature extraction layer. Moreover, the second best features have been generated using 60×60 sized patches (5th feature extraction layer). The classification accuracies of the 4th and 5th feature extraction layers are 90.69% and 90.67% respectively.
- The worst one is the 1st feature extraction layer since the features of this layer yielded 88.87% classification accuracy and this layer has used 60×60 sized patches.

- We have compared the commonly used shallow classifiers and the best resulting classifier is the SVM classifier. Therefore, we have used this classifier.
- Three validations were used to attain the classification results.

5.2.2 Advantages

- Our team has found that Swin-LBP can improve LBP's classification performance by approximately 11%.
- We tested the recommended Swin-LBP approach on a large dataset of 12,230 urine cell images, achieving an impressive classification accuracy of 92.60%. This result highlights the potential of handcrafted models to achieve outstanding performance on large image datasets.
- Our study also revealed that Swin-LBP outperforms published deep models developed on the same urine sediment image dataset.
- The Swin-LBP model proposed in our study is simple and can be easily implemented by researchers to address image classification tasks.

5.2.3 Limitations

- Although fine-tuning operations can achieve higher classification performance, we require a fast-responding model. As such, we opted not to use any optimization techniques.
- To further validate the proposed Swin-LBP model's classification performance, additional urine image datasets could be used. Such datasets could provide more comprehensive insights into the model's generalization capabilities and its potential to address real-world image classification problems.

6 Conclusions

In this work, we proposed a computationally lightweight yet accurate model for automated analysis of urine sediments using the Swin Transformer architecture with shifted windows-based patch division. Our approach enabled global and local textural feature extraction, selection, and classification using LBP, NCA, and SVM, respectively. The model achieved excellent results on a derived 7-class study dataset comprising 12,330 urine images, with a classification accuracy of 92.60%. Our model has low time complexity and is simple yet accurate, making it suitable for real-world urine sediment analysis.

Our study also highlights the versatility and utility of computer vision-inspired shifted windows-based patch division for general image classification problems that enable multilevel downstream feature extraction using handcrafted feature engineering. Specifically, our results confirm the feasibility of the Swin-LBP approach in biomedical image analysis applications.

Our study presents a novel method for the automated analysis of urine sediments that achieves excellent classification accuracy with a computationally efficient approach. Additionally, our use of shifted windows-based patch division provides a promising technique for general image classification problems that can enable multilevel downstream feature extraction using handcrafted feature engineering.

In future work, we plan to propose automated urine cell counting and classification applications where the Swin architecture can be combined with other image classification options such as transfer learning, further extending the utility of our approach.

Declarations

Conflict of interest The authors of this manuscript declare no conflict of interest.

References

1. Tasoglu S (2022) Toilet-based continuous health monitoring using urine. Nat Rev Urol 19:219–230
2. Suhail K, Brindha D (2021) A review on various methods for recognition of urine particles using digital microscopic images of urine sediments. Biomed Signal Process Control 68:102806
3. Jiménez-Zucchet N, Alejandro-Zayas T, Alvarado-Macedo CA, Arreola-Illescas MR, Benítez-Araiza L, Bustamante-Tello L, Cruz-Martínes D, Falcón-Robles N, Garduño-González L, López-

Romahn MC (2019) Baseline urinalysis values in common bottlenose dolphins under human care in the Caribbean. J Vet Diagn Invest 31:426–433

4. Li Q, Yu Z, Qi T, Zheng L, Qi S, He Z, Li S, Guan H (2020) Inspection of visible components in urine based on deep learning. Med Phys 47:2937–2949

5. Cho J, Oh KJ, Jeon BC, Lee S-G, Kim J-H (2019) Comparison of five automated urine sediment analyzers with manual microscopy for accurate identification of urine sediment. Clin Chem Lab Med (CCLM) 57:1744–1753

6. Laiwejpithaya S, Wongkrajang P, Reesukumal K, Bucha C, Meepanya S, Pattanavin C, Khejonnit V, Chuntarut A (2018) UriSed 3 and UX-2000 automated urine sediment analyzers vs manual microscopic method: a comparative performance analysis. J Clin Lab Anal 32:e22249

7. Khalid ZM, Hawezi RS, Amin SRM (2022) Urine sediment analysis by using convolution neural network. In: 2022 8th International Engineering Conference on Sustainable Technology and Development (IEC), pp 173–178, IEEE

8. Liu H, Li Q, Zhang Y, Huang D, Yu F (2022) Consistency analysis of the Sysmex UF-5000 and Atellica UAS 800 urine sedimentation analyzers. J Clin Lab Anal 36:e24659

9. Li T, Jin D, Du C, Cao X, Chen H, Yan J, Chen N, Chen Z, Feng Z, Liu S (2020) The image-based analysis and classification of urine sediments using a LeNet-5 neural network. Comput Methods Biomech Biomed Eng: Imaging Vis 8:109–114

10. Houssein EH, Helmy BE-D, Oliva D, Jangir P, Premkumar M, Elngar AA, Shaban H (2022) An efficient multi-thresholding based COVID-19 CT images segmentation approach using an improved equilibrium optimizer. Biomed Signal Process Control 73:103401

11. Devi RM, Premkumar M, Jangir P, Kumar BS, Alrowaili D, Nisar KS (2022) BHGSO: binary hunger games search optimization algorithm for feature selection problem. CMC-Comput Mater Continua 70:557–579

12. Premkumar M, Jangir P, Sowmya R, Elavarasan RM (2021) Many-objective gradient-based optimizer to solve optimal power flow problems: analysis and validations. Eng Appl Artif Intell 106:104479

13. Premkumar M, Sowmya R, Umashankar S, Jangir P (2021) Extraction of uncertain parameters of single-diode photovoltaic module using hybrid particle swarm optimization and grey wolf optimization algorithm. Mater Today: Proc 46:5315–5321

14. Liang Y, Tang Z, Yan M, Liu J (2018) Object detection based on deep learning for urine sediment examination. Biocybern Biomed Eng 38:661–670

15. Zhang X, Jiang L, Yang D, Yan J, Lu X (2019) Urine sediment recognition method based on multi-view deep residual learning in microscopic image. J Med Syst 43:1–10

16. Pan J, Jiang C, Zhu C (2018) Classification of urine sediment based on convolution neural network. In: AIP Conference Proceedings, AIP Publishing LLC, pp 040176

17. Ji Q, Li X, Qu Z, Dai C (2019) Research on urine sediment images recognition based on deep learning. IEEE Access 7:166711–166720

18. Velasco JS, Cabatuan MK, Dadios EP (2019) Urine sediment classification using deep learning. Lect Notes Adv Res Electr Electron Eng Technol, :180–185

19. Liu W, Li W, Gong W (2020) Ensemble of fine-tuned convolutional neural networks for urine sediment microscopic image classification. IET Comput Vision 14:18–25

20. Khan AA, Laghari AA, Awan SA (2021) Machine learning in computer vision: a review. EAI Trans Scalable Inf Syst 8:e4

21. Hossain MS, Bilbao J, Tobón DP, Muhammad G, Saddik AE (2022) Special issue deep learning for multimedia healthcare. Multimed Syst 28(4):1147–1150

22. Chu H, He Z, Liu S, Liu C, Yang J, Wang F (2022) Deep neural network for point sets based on local feature integration. Sensors 22:3209

23. Wang L, Fang S, Li R, Meng X (2022) Building extraction with vision transformer. IEEE Trans Geosci Remote Sens 60:1–11

24. Wang F, Rao Y, Luo Q, Jin X, Jiang Z, Zhang W, Li S (2022) Practical cucumber leaf disease recognition using improved Swin Transformer and small sample size. Comput Electron Agric 199:107163

25. Xu P, Zhu X, Clifton DA (2022) Multimodal learning with transformers: a survey. arXiv preprint arXiv:2206.06488

26. Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S (2020) An image is worth 16 × 16 words: transformers for image recognition at scale. arXiv preprint arXiv:2010.11929

27. Liu Z, Lin Y, Cao Y, Hu H, Wei Y, Zhang Z, Lin S, Guo B (2021) Swin transformer: hierarchical vision transformer using shifted windows. In: Proceedings of the IEEE/CVF international conference on computer vision, pp 10012–10022

28. Ojala T, Pietikainen M, Maenpaa T (2002) Multiresolution grayscale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell 24:971–987

29. Yang W, Wang K, Zuo W (2012) Neighborhood component feature selection for high-dimensional data. J Comput 7:161–168

30. Peterson LE (2009) K-nearest neighbor. Scholarpedia 4:1883

31. Liang Y, Kang R, Lian C, Mao Y (2018) An end-to-end system for automatic urinary particle recognition with convolutional neural network. J Med Syst 42:1–14

32. Yan M, Liu Q, Yin Z, Wang D, Liang Y (2020) A bidirectional context propagation network for urine sediment particle detection in microscopic images. In: ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp 981–985, IEEE

33. Powers DM (2020) Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. arXiv preprint arXiv:2010.16061

34. Warrens MJ (2008) On the equivalence of Cohen's kappa and the Hubert-Arabie adjusted Rand index. J Classif 25:177–183

35. Kang R, Liang Y, Lian C, Mao Y (2018) CNN-based automatic urinary particles recognition. arXiv preprint arXiv:1803.02699

Authors and Affiliations

Mehmet Erten[1] · Prabal Datta Barua[2] (iD) · Ilknur Tuncer[3] · Sengul Dogan[4] · Mehmet Baygin[5] ·
Turker Tuncer[4] · Ru-San Tan[6,7] · U. Rajendra Acharya[8]

✉ Prabal Datta Barua
  Prabal.Barua@usq.edu.au

  Mehmet Erten
  mehmeter23@gmail.com

  Ilknur Tuncer
  ilknur.tuncer@icisleri.gov.tr

  Sengul Dogan
  sdogan@firat.edu.tr

  Mehmet Baygin
  mehmet.baygin@erzurum.edu.tr

  Turker Tuncer
  turkertuncer@firat.edu.tr

  Ru-San Tan
  tanrsnhc@gmail.com

  U. Rajendra Acharya
  Rajendra.Acharya@usq.edu.au

[1] Department of Biochemistry, Elazig Fethi Sekin City
     Hospital, Elazig, Turkey

[2] School of Business (Information System), University of
     Southern Queensland, Toowoomba, Australia

[3] Elazig Governorship, Interior Ministry, Elazig, Turkey

[4] Department of Digital Forensics Engineering, College of
     Technology, Firat University, 23119 Elazig, Turkey

[5] Department of Computer Engineering, College of
     Engineering, Erzurum Technical University, Erzurum,
     Turkey

[6] Department of Cardiology, National Heart Centre Singapore,
     Singapore, Singapore

[7] Duke-NUS Medical School, Singapore, Singapore

[8] School of Mathematics, Physics and Computing, University
     of Southern Queensland, Springfield, Australia