Frontiers in Health Informatics
ISSN-Online: 2676-7104
www.healthinformaticsjournal.com
2024; Vol 13: Issue 3
Open Access

Development of Infection Category Detection in Diabetic Foot Ulcers Using Machine Learning

Haryanto[1]*, Annisa Rahmawati[1], Yuyun Nisaul Khairillah[2], Indri Erwhani[1], Ditha Astuti Purnamawati[1], Tiara Amalia Mentari[3], Erlis Eka Fitriana[3]

1 Department of Nursing, Institut Teknologi dan Kesehatan Muhammadiyah Kalimantan Barat, Pontianak, Indonesia

2 Department of Biotechnology, Institut Teknologi dan Kesehatan Muhammadiyah Kalimantan Barat, Pontianak, Indonesia

3 Student of Magister Nursing, Institut Teknologi dan Kesehatan Muhammadiyah Kalimantan Barat, Pontianak, Indonesia

Cite this paper as: Haryanto, Annisa Rahmawati, Yuyun Nisaul Khairillah, Indri Erwhani, Ditha Astuti Purnamawati, Tiara Amalia Mentari, Erlis Eka Fitriana (2024) Development of Infection Category Detection in Diabetic Foot Ulcers Using Machine Learning. *Frontiers in Health Informatics, 13* (3), 1669-1675

A B S T R A C T

Introduction: *Traditional infection detection methods, including clinical examinations and microbiological cultures, are time-consuming. Artificial Intelligence (AI), particularly machine learning and deep learning, offers a promising alternative by swiftly and accurately analyzing medical data.*

Objectives: *This study aimed to identify the role of AI in detecting infections in DFUs using image recognition algorithms and prediction models based on patient clinical parameters.*

Methods: *An image dataset of DFU infections, categorized into three classes (mild, moderate, and severe), was utilized, consisting of 90 images (30 images per class). Features were extracted using the Gray Level Co-occurrence Matrix (GLCM) from segmented images. To handle non-linear data, the Support Vector Machine (SVM) algorithm was employed with a kernel trick, mapping lower-dimensional data to higher dimensions to achieve linear separability. The model's performance was evaluated using accuracy, precision, and recall.*

Results: *The SVM classifier achieved an overall accuracy of 71.4%. It performed exceptionally for moderate infections, achieving perfect precision, recall, and F1-score of 1.00. For mild infections, the precision was 0.50, recall was 1.00, and F1-score was 0.67, indicating some misclassifications. Severe infections had a precision of 1.00, recall of 0.67, and F1-score of 0.80, suggesting a conservative approach that missed some severe cases.*

Conclusions*: Future improvements could include hyperparameter tuning, expanding training datasets, and employing advanced techniques like SMOTE or ADASYN for class imbalance. Despite current limitations, AI shows significant potential in revolutionizing diabetic wound management, offering more timely and effective infection detection, thus improving clinical outcomes.*

Keywords: *foot ulcer wound infectionartificial intelligencemachine learningsupport vector machine*

Introduction

Diabetes mellitus is a chronic metabolic disease characterized by prolonged hyperglycemia due to disruptions in insulin secretion, insulin action, or both [1]. One of the serious complications frequently seen in diabetes patients is diabetic foot ulcer, an open wound typically occurring on the feet of diabetic individuals [2]. Infection in diabetic foot ulcers poses a major clinical challenge as it can worsen patient conditions and increase the risk of amputation [2].

Managing infection in diabetic foot ulcers requires timely and accurate diagnosis to ensure appropriate treatment. Conventional methods for detecting infection involve clinical examination and microbiological culture [3], which often take time and may not yield rapid results. This is where Artificial Intelligence (AI) technology can play a crucial role. AI has the potential to analyze medical data swiftly and accurately, offering more efficient solutions for early detection of infections in diabetic foot ulcers[4,5].

AI technology, particularly machine learning and deep learning, has shown remarkable capabilities in analyzing medical images and clinical [6–8]. By utilizing algorithms trained on large datasets, AI can identify patterns associated with infection faster than traditional methods[9]. Several studies have demonstrated that AI can detect infections with high accuracy, aiding clinicians in making faster and more precise clinical decisions [7–9].

In the context of detecting infections in diabetic foot ulcers, AI applications can vary from analyzing wound images to predicting infection risks based on patient clinical parameters. For instance, image recognition algorithms can analyze wound photos and identify signs of infection such as color changes, exudate, and inflammatory signs[10–12]. Additionally, AI-based prediction models can leverage patient historical data such as blood glucose levels, blood pressure, and medical history to estimate the likelihood of infection [13].

The development of AI technology offers substantial potential to enhance early detection of infections in diabetic wounds. AI can be utilized to analyze medical images, clinical data, and other parameters swiftly and accurately. Thus, this technology has the potential to revolutionize the paradigm of diabetic wound management, enabling more timely and effective interventions.

While the potential of AI in detecting infections in diabetic foot ulcers is promising, there are several challenges to overcome. These include the need for high-quality data to train algorithms, integration of AI technology into clinical practices, and ensuring the security and privacy of patient data. By addressing these challenges, AI can become a valuable tool in improving diabetic ulcer management and reducing infection-related complications. The purpose of this study is to identify the use of AI technology particularly SVM (Support Vector Machine) in detecting infections in diabetic foot ulcers. The purpose of this study is to identify the use of AI technology particularly SVM (Support Vector Machine) in detecting infections in diabetic foot ulcers.

MATERIAL AND Methods

This study will employ an experimental approach where the outcomes of AI technology in detecting infections in diabetic wounds. The research methodology is presented in Fig 1 including:

Image Data Set

The dataset of external wound images was sourced from the PKU Muhammadiyah Klinik Kitamura, Pontianak, Indonesia. Figure 2 displays the original image with a resolution of 640x640 pixels. The image dataset of DFU Infection is categorized into three classes: mild, moderate, and severe. Each class comprises 30 images, making a total of 90 images in the dataset.

b. The features utilized in this research are derived from the results of image segmentation.

These features are normalized within the range [0,1] and then used as input data for the SVM classifier. The features include average, variance, standard deviation, and skewness. Feature extraction is employed to transform image data from a matrix format to a vector format. This research uses Gray Level Co-occurrence Matrix (GLCM) for feature extraction, a technique for image texture analysis that captures the relationship between two adjacent pixels in terms of intensity, distance, and grayscale angle. GLCM considers eight angles: 0º, 45º, 90º, 135º, 180º, 225º, 270º, and 315º. The distance parameter in GLCM calculations is based on the number of pixels between reference and neighboring pixels. Four features extracted from GLCM are homogeneity, contrast, correlation, and energy. Entropy provides information about the randomness of the image. Homogeneity measures the closeness of the distribution of elements in the GLCM to the diagonal, with a value of 1 indicating perfect diagonal homogeneity and a value of 0 indicating perfect vertical homogeneity.

Contrast measures the intensity variation between pixels, highlighting differences in high and low colors within the image. These values are crucial for analyzing and determining the type of infection.

c. Classification Using Support Vector Machine

Datasets with two classes that can be separated by a straight line are known as linear data, while datasets that cannot be separated in this way are referred to as non-linear data. Non-linear data typically require separation by irregular curves. The SVM algorithm struggles to separate non-linear data, so a kernel trick is used to map non-linear data from lower dimensions to higher dimensions, making it possible to separate them as linear data.

d. Model Evaluation Classification

Calculating accuracy in classification is the crucial final step, as it determines whether the algorithm used is appropriate for the given problem. This step will assess the effectiveness of the SVM algorithm for classifying external wound images. The model will be evaluated using metrics such as accuracy, precision, recall and F1 score as shown in Eq. 1,2, 3 and 4.

$$\text{Accuracy} = [TN + TP] \div [TN + FP + FN + TP] \quad (1)$$

$$\text{Precision} = [TP] \div [TP + FP] \quad (2)$$

$$\text{Recall} = [TP] \div [TP + FN] \quad (3)$$

$$\text{F1 score} = 2 \times [\text{Recall} \times \text{Precision}] \div [\text{Recall} + \text{Precision}] \quad (4)$$

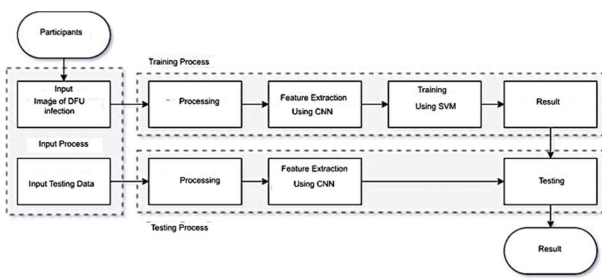

Fig 1: SVM Classification Flowchart

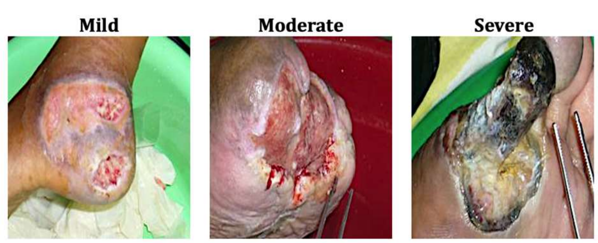

Fig 2: Wound Data Set

Results

Precision measures the accuracy of the model in classifying positive instances. For the moderate class, precision is 1.00, indicating all predictions classified as moderate are correct. For the mild class, precision is 0.50, indicating only half of the predictions are correctly classified as mild. For the severe class, precision is 1.00, indicating all predictions classified as severe are correct.

Recall (or sensitivity) measures how well the model can identify all positive instances. For the moderate and mild classes, recall is 1.00, indicating the model identifies all instances correctly for these classes. However, for the severe class, recall is 0.67, indicating the model misses some severe instances.

F1-score is the harmonic mean of precision and recall, providing an overall measure of model performance for each class. The lowest F1-score is observed for the mild class at 0.67, suggesting a less balanced precision-recall trade-off compared to other classes. The other two classes exhibit higher F1-scores (1.00 for moderate and 0.80 for severe). In summary, despite a relatively high accuracy (0.71), class-specific evaluations reveal varying performance depending on the severity level of the classified conditions (Tab 1).

Table 1. Actual values and Predicted values (Confusion Matrix)

	precision	recall	f1-score	support
Moderate infection	1.00	1.00	1.00	45
Mild infection	0.50	1.00	0.67	15
Severe infection	1.00	0.67	0.80	30
accuracy			0.71	90
macro avg	0.83	0.89	0.82	90
weighted	0.86	0.71	0.75	90

Discussion

This is the first study to identify and classify infections in foot ulcers using machine learning particulary SVM. The aim of this study is to identify severity of infection from image of the diabetic foot using machine learning particularly SVM. The SVM classifier demonstrated a reasonable performance in classifying diabetic wound infections, achieving an overall accuracy of 71.4%.

In a similar study [14] to ours, The study focuses on the early identification of risk factors associated with the development of diabetic foot ulcer (DFU) using machine learning techniques. It aims to find the association between various clinical and biochemical risk factors with DFU and develop a prediction model using different machine learning algorithms. The observational study included 80 type 2 diabetes mellitus (T2DM) patients with DFU and 80 T2DM patients without DFU. Clinical and laboratory data were analysed using algorithms such as Support Vector Machines (SVM-Poly K), Naive Bayes (NB), K-nearest neighbor (KNN), Random Forest (RF), and three ensemble learners: Stacking C, Bagging, and AdaBoost. These algorithms were used to construct prediction models for discriminating between the two groups (stage I classification) and ulcer type classification (stage II classification). Ensemble learning outperformed individual classifiers in various performance evaluation metrics. According to the results, SVM-poly K has accuracy and specificity were 93.8% and 93.8% respectively.

Similarly to our, a study [15] suggested Convolutional Neural Network (CNN) model showed high accuracy and effectiveness in identifying infections (91.9%). However, our study used SVM. CNNs is deep learning algorithms designed to process input images by convolving them with filters or kernels to extract features. In this process, an NxN image is convolved with an fxf filter, where the convolution operation learns features uniformly across the entire image [16]. Meanwhile, SVM is a machine learning method used for classification and regression tasks. SVM works by finding an optimal hyperplane that separates data into different classes with maximum margin. This method is effective for high-dimensional datasets and can handle both linear and nonlinear data using suitable kernel functions. SVM is well-known for its strong generalization capability and ability to mitigate overfitting, particularly on relatively small to medium-sized datasets [17]. Therefore, the choice

between SVM and CNN depends on the type of data, the size of the dataset, and the available computational resources.

In our research, the classifier exhibited excellent performance in identifying moderate wounds, with perfect precision, recall, and F1-score of 1.00. This indicates that all instances of moderate wounds were correctly classified without any false positives or false negatives. For mild wounds, the classifier showed a precision of 0.50 and a recall of 1.00, resulting in an F1-score of 0.67. This suggests that while the classifier successfully identified all actual mild wounds (high recall), it also misclassified some instances as mild that were not mild (low precision). This indicates a need for improving the classifier's ability to distinguish mild wounds from other categories. In the case of severe wounds, the classifier achieved a precision of 1.00 and a recall of 0.67, resulting in an F1-score of 0.80. The high precision indicates that all predicted severe wounds were indeed severe, but the lower recall implies that the classifier missed some actual severe wound instances. This highlights the classifier's conservative approach in predicting severe wounds, which may result in underreporting of this category. The macro average metrics provide an unweighted mean across all classes, reflecting an overall balanced performance. Therefore, the classifier performs exceptionally well in identifying moderate wounds but requires improvement in distinguishing mild and severe wounds. Enhancements in precision are needed for mild wounds, and improvements in recall are necessary for severe wounds to mitigate potential underreporting. This statement aligns with findings from research published by Burges and Schölkopf [18].

The weighted average metrics take into account the support (number of instances) for each class, providing a more comprehensive view of the classifier's performance. The differences between the macro and weighted averages suggest that the distribution of classes affects the overall performance metrics, with the classifier performing better on classes with more instances. Therefore, weighted average metrics consider the support (number of instances) for each class, offering a more comprehensive evaluation of the classifier's performance. The disparities between macro and weighted averages indicate that the class distribution significantly influences overall performance metrics, with the classifier demonstrating stronger performance on classes that have more instances. This statement is supported by Weiss and Provost [19],

that how classifier performance is influenced by class distribution can assist practitioners in selecting training data.

To improve the classifier's performance, particularly for the mild and severe categories, further strategies could include additional hyperparameter tuning, increasing the size and diversity of the training dataset, and employing more advanced techniques for handling class imbalance, such as Synthetic Minority Over-sampling Technique (SMOTE) or Adaptive Synthetic (ADASYN) sampling.

conclusion

In conclusion, while the SVM classifier shows promising results, especially for Moderate wounds, there is room for improvement in classifying Mild and Severe wounds. Future work should focus on addressing these areas to enhance the classifier's accuracy and reliability in clinical applications.

ACKNOWLEDGMENT

The authors thank to Majelis Pendidikan Tinggi Penelitian dan Pengembangan Pimpinan Pusat Muhammadiyah and Pusat Penelitian Pengabdian Masyarakat Institut Teknologi dan Kesehatan Muhammadiyah Kalimantan Barat for supporting this study.

AUTHOR’S CONTRIBUTION

HY and AR: conceptualization, methodology, software; SK, YNK, and IE: data curation, original draft preparation; DAP, TAM, and EEF: visualization, investigation; HY and YNK: supervision. All authors read and approved the final manuscript.

CONFLICTS OF INTEREST

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

FINANCIAL DISCLOSURE

This study was supported by Majelis Pendidikan Tinggi Penelitian dan Pengembangan Pimpinan Pusat Muhammadiyah (RISETMU scheme) (Grant Number: 0258.476/I.3/D/2024).

ETHICS APPROVAL

Ethical approval for this study was obtained from the ethical committee of the Institute of Technology and Health Muhammadiyah West Kalimantan’s ethics committee board. (Number: No. 138/II.I.AU/KET.ETIK/III/2024)

Haryanto, ID ORCID: https://orcid.org/0000-0002-3711-1885

Refrences

1. Galicia-garcia U, Benito-vicente A, Jebari S, Larrea-sebal A. Costus ignus: Insulin plant and it’s preparations as remedial approach for diabetes mellitus. Int J Mol Sci. 2020;1–34.

2. Stancu B, Ilyés T, Farcas M, Coman HF, Chiș BA, Andercou OA. Diabetic Foot Complications: A Retrospective Cohort Study. Int J Environ Res Public Health. 2023;20(1).

3. Au/Australianprescriber O. Full text free online at nps Non-culture methods for detecting infection. 2016;39(5):171–5. Available from: http://dx.doi.org/10.18773/

4. Mousa KM, Mousa FA, Mohamed HS, Elsawy MM. Prediction of Foot Ulcers Using Artificial Intelligence for Diabetic Patients at Cairo University Hospital, Egypt. SAGE Open Nurs. 2023;9.

5. Chan KS, Meng Y, Audrey C, Min H, Rhan L, Chong C, Zhang L, Wei G, Tan L, Chandrasekar S, Joseph Z. Clinical validation of an artificial intelligence-enabled wound imaging mobile application in diabetic foot ulcers. 2022;(April 2021):114–24.

6. Pinto-Coelho L. How Artificial Intelligence Is Shaping Medical Imaging Technology: A Survey of Innovations and Applications. Bioengineering. 2023;10(12).

7. Karalis VD. The Integration of Artificial Intelligence into Clinical Practice. Appl Biosci. 2024;3(1):14–44.

8. Krishnan G, Singh S, Pathania M, Gosavi S, Abhishek S, Parchani A, Dhar M. Artificial intelligence in clinical medicine: catalyzing a sustainable global healthcare paradigm. Front Artif Intell. 2023;6.

9. Shelke YP, Badge AK, Bankar NJ. Applications of Artificial Intelligence in Microbial Diagnosis. Cureus. 2023;15(11):1–8.

10. Chairat S, Chaichulee S, Dissaneewate T, Wangkulangkul P. AI-Assisted Assessment of Wound Tissue with Automatic Color and Measurement Calibration on Images Taken with a Smartphone. 2023;1–22.

11. Curti N, Merli Y, Zengarini C, Starace M, Rapparini L, Marcelli E, Carlini G, Buschi D, Castellani GC, Maria B, Bianchi T, Giampieri E. Automated Prediction of Photographic Wound Assessment Tool in Chronic Wound Images. J Med Syst [Internet]. 2024; Available from: https://doi.org/10.1007/s10916-023-02029-9

12. Howell RS, Liu HH, Khan AA, Woods JS, Lin LJ, Saxena M, Saxena H, Castellano M. Development of a Method for Clinical Evaluation of Artificial Intelligence – Based Digital Wound Assessment Tools. 2021;1–12.

13. Sampa MB, Biswas T, Rahman MS, Aziz NHBA, Hossain MN, Aziz NAA. A Machine Learning Web App to Predict Diabetic Blood Glucose Based on a Basic Noninvasive Health Checkup, Sociodemographic Characteristics, and Dietary Information: Case Study. JMIR Diabetes. 2023;8(1).

14. Nanda R, Nath A, Patel S, Mohapatra E. Machine learning algorithm to evaluate risk factors of diabetic foot ulcers and its severity. Med Biol Eng Comput [Internet]. 2022;60(8):2349–57. Available from: https://doi.org/10.1007/s11517-022-02617-w

15. Yogapriya J, Chandran V, Sumithra MG, Elakkiya B, Shamila Ebenezer A, Suresh Gnana Dhas C. Automated Detection of Infection in Diabetic Foot Ulcer Images Using Convolutional Neural Network. J Healthc Eng. 2022;2022.

16. Zeiler MD, Fergus R. Visualizing and understanding convolutional networks. Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics). 2014;8689 LNCS(PART 1):818–33.

17. Jun Z. The Development and Application of Support Vector Machine. J Phys Conf Ser. 2021;1748(5).

18. Burges CJC, Schölkopf B. Improving the accuracy and speed of support vector machine. Adv Neural Inf Process Syst. 1997;(July 1997):375–81.

19. Weiss GM, Provost F. The Effect of Class Distribution on Classifier Learning : An Empirical Study The Effect of Class Distribution on Classifier Learning : An Empirical Study. 2013;(September 2001).