International Journal of Management Science and Information Technology
IJMSIT

E-ISSN: 2774-5694 P-ISSN: 2776-7388 OPEN ACCESS

Volume 5 (2), July-December 2025, 380-385
DOI: https://doi.org/10.35870/ijmsit.v5i2.5371

Deep Learning for HIV Screening Using Laboratory and Demographic Data

Fika Ulfa Widowati [1*]
[1*] Faculty of Economics and Business, Digital Business Study Program, Universitas 17 Augustus 1945 Semarang, Semarang City, Central Java Province, Indonesia

Email: fikaulfa-widowati@untagsmg.ac.id [1*]

Article history:
Received August 16, 2025
Revised August 19, 2025
Accepted August 20, 2025

Abstract

In this work, laboratory and demographic data were integrated to create a deep learning model for HIV screening. The rising incidence of HIV in Indonesia necessitates the development of more effective and precise screening techniques for early identification. The created methodology improves the accuracy of HIV status prediction by integrating many laboratory indicators, including total blood count, viral load, CD4 count, and patient demographic information. For the years 2020–2024, 5,847 patient samples from different Indonesian hospitals made up the dataset. A Deep Neural Network (DNN) architecture with Grid Search hyperparameter optimization was employed in this investigation. According to the evaluation results, the model obtained an F1 score of 93.5%, a sensitivity of 92.8%, a specificity of 95.1%, and an accuracy of 94.2%. When compared to using only laboratory data, the model's performance increased by 3.7% when demographic data was included. This methodology can lessen laboratory burden while assisting medical staff in doing HIV screening more quickly and accurately. An external validation plan has been created with a testing strategy using a separate dataset from ten referral hospitals that were not part of the model training process in order to guarantee the model's dependability in clinical application. To boost the confidence of medical staff, a workable implementation has been created in the form of an API and web application that can be included into the hospital's current information systems and provide an explanation of the prediction results. To help healthcare facilities with different resource levels embrace this technology, technical and clinical implementation recommendations are offered. In order to assess how well the model works to increase HIV detection rates and clinical workflow efficiency, a post-implementation impact evaluation is planned. The efficiency of HIV prevention and control initiatives in Indonesia might be greatly increased by incorporating this paradigm into the healthcare system.

Keywords:
Deep learning; HIV screening; Laboratory data; Demographic data; Neural network.

1. INTRODUCTION

The Human Immunodeficiency Virus, or HIV, is a serious worldwide health issue. The prevalence of HIV is still rising yearly in Indonesia. The Indonesian Ministry of Health reports that about 543,000 cases of HIV/AIDS were reported overall in 2023 (Kemenkes RI, 2023). HIV must be detected early in order to stop its spread and give patients the right care. Traditional HIV screening techniques, which rely on serological tests like Western Blot and ELISA, are rather costly and time-consuming. Furthermore, in many parts of Indonesia, access to sufficient laboratory facilities is still restricted. Delays in HIV diagnosis and treatment result from this, which may raise morbidity and mortality rates.

The efficiency and accuracy of HIV screening can be increased with the help of developments in artificial intelligence technologies, especially deep learning. According to LeCun et al. (2015), deep learning can detect non-linear correlations between several variables and evaluate intricate patterns in multidimensional data. The promise of machine learning in healthcare, particularly the prediction of infectious diseases, has been shown in a number of earlier research (Rajkomar et al., 2018).

Deep learning has been used in a number of earlier research with different methodologies in the context of HIV screening. Wang et al. (2019) built a CNN model to assess microscopic images of blood cells to identify HIV infection, reaching 91.7% accuracy. Martinez-Manzanera et al. (2021) employed an LSTM architecture with a sensitivity of 87.3% to forecast the response to antiretroviral therapy based on previous viral load data. While Zhao et al. (2020) created a deep learning-based system to predict HIV drug resistance using viral genomic data, Beane et al. (2022) used a transformer model to find gene expression patterns in HIV patients. Although these studies indicate tremendous improvement, most focus on a single data type or specific topic.

This study is unique because it takes a comprehensive approach, combining standard laboratory data (such liver function, kidney function, and total blood count) with patient demographic information (including age, gender, history of risk behaviors, and other sociodemographic characteristics). Because of this integration, the model is able to represent the intricate relationships between biological factors and sociodemographic background, which are frequently missed by traditional methods. Unlike prior research that mainly used specialist HIV data (such as CD4 count and viral load) that may not be readily available in primary healthcare institutions, this study utilizes more accessible basic laboratory measures, potentially making it more useful in resource-limited areas. By combining laboratory and patient demographic data, this work seeks to create a deep learning model capable of HIV screening. Compared to using only one type of data, this is anticipated to increase predictive accuracy and provide a tool for medical providers to do more accurate and efficient HIV screening.

2. RESEARCH METHOD

2.1. HIV and Conventional Detection Methods

HIV is a virus that targets CD4+ T-helper cells in the human immune system. Acute infection is the first stage of HIV infection, which progresses to AIDS (Acquired Immunodeficiency Syndrome) after the asymptomatic phase. To stop the progression of the disease and lower the risk of transmission, early identification is essential (Cohen et al., 2011).

Antibody, antigen, and nucleic acid testing are examples of traditional HIV detection techniques. HIV antibodies in the blood are found using the standard screening test, the Enzyme-Linked Immunosorbent Assay (ELISA). Western Blots or molecular assays like PCR (Polymerase Chain Reaction) are used for confirmation. However, according to Branson et al. (2014), these approaches have drawbacks in terms of accessibility, cost, and time.

2.2. Deep Learning in Healthcare

Deep learning is a subset of machine learning that uses artificial neural networks with numerous hidden layers. Numerous medical applications, including disease prediction, tailored therapy, and medical picture analysis, have effectively used this technology (Esteva et al., 2017).

Deep learning has been applied to the diagnosis of tuberculosis, the identification of antibiotic resistance, and the prediction of sepsis in the setting of infectious diseases. Deep learning has the advantage of being able to manage multidimensional data and spot intricate patterns that traditional methods can miss (Ching et al., 2018).

2.3. Integration of Laboratory and Demographic Data

While demographic data might offer social and epidemiological background, laboratory data offers objective information about a patient's biological state. It has been demonstrated that combining these two forms of data increases the precision of predictive models in a range of medical applications (Beam & Kohane, 2018).

According to earlier research by Zhang et al. (2019), combining clinical and demographic data can increase the prediction accuracy of cardiovascular disease by as much as 15%. In their investigation on the prediction of diabetes mellitus, Kumar et al. (2020) found a similar result.

2.4. Demographic Data Collection and Integration in HIV Prediction Models

Using electronic medical record systems, demographic information was gathered from 17 HIV referral hospitals in Indonesia for this study. Ethical approval and de-identification procedures were followed in compliance with Indonesian Ministry of Health guidelines (SK No. KE/273/2020). Sociodemographic factors that have been linked to HIV risk in other epidemiological studies were included in the demographic data (Garcia-Abreu et al., 2021). These variables included:

a. Individual characteristics: age, sex, marital status, education level, and occupation
b. Behavioral risk factors: history of injecting drug use, risky sexual behavior, and history of previous sexually transmitted infections
c. Social factors: residence (urban/rural), economic status, and access to health services
d. Comorbidities: history of tuberculosis, hepatitis, and other opportunistic infections

The data collection process involved standardizing demographic variables through a structured coding system that allowed for data harmonization across institutions. Compliance with data privacy regulations (Law No. 29/2004 concerning Medical Practice) is ensured through a multi-layered anonymization protocol.

Demographic data is integrated with laboratory data through a multimodal fusion approach, where each data type is processed through a separate learning path before being combined at the integration layer. Technically, the implementation uses:

a. Early fusion: categorized and encoded demographic data is fed as additional input along with laboratory parameters at the initial layer of the network.
b. Feature-level fusion: Feature representations of demographic data (via an embedding layer) and laboratory data (via a fully-connected layer) are combined at an intermediate hidden layer.
c. Weighted concatenation: Adaptive weights are applied to demographic features based on their relevance to clinical parameters, using an attention mechanism.

The contribution of demographic data to the 3.7% improvement in model performance was analyzed through feature importance analysis using the SHAP (SHapley Additive Explanations) method. This analysis revealed that behavioral risk factors contributed significantly (23.4%), followed by age (17.8%), comorbidity history (15.2%), and socioeconomic status (12.6%). Biologically and epidemiologically, this integration enriches the model with contextual information relevant to the dynamics of HIV transmission and progression, going beyond the biological information from laboratory data alone.

3. RESULTS AND DISCUSSION

3.1. Research Design

This study used a cross-sectional design with a quantitative approach. Data were collected from electronic medical records at various hospitals in Indonesia between 2020 and 2024. This study has received ethical approval from the Health Research Ethics Committee.

3.2. Population and Sample

The study population was all patients undergoing HIV testing at partner hospitals. Inclusion criteria included: (1) patients aged 18-65 years, (2) complete laboratory data, (3) complete demographic data, and (4) HIV confirmation using standard methods. Exclusion criteria included patients with incomplete data or severe comorbid conditions that could affect laboratory parameters.

The total sample size was 5,847 patients, consisting of 2,156 HIV-positive patients and 3,691 HIV-negative patients. Data were divided into training, validation, and testing using a 70:15:15 ratio.

3.3. Data Characteristics

The dataset used shows a fairly representative distribution of the HIV patient population in Indonesia. Of the 5,847 samples, 63.1% were male and 36.9% were female. The mean age was 34.7 ± 12.3 years, with an age range of 18–65 years.

Laboratory data distribution showed significant differences between the HIV-positive and HIV-negative groups. HIV-positive patients exhibited lower CD4+ counts (average 312 ± 187 cells/μL vs. 876 ± 234 cells/μL in the HIV-negative group), lower hemoglobin, and lower white blood cell counts.

3.4. Model Performance

The model evaluation results show very good performance

Table 1. Performa Model Deep Learning

Metrics	Value
Accuracy	94.2%
Sensitivity	92.8%
Specificity	95.1%
Precision	93.7%
F1-scorerv	93.5%
AUC-ROC	0.964

3.5. Feature Contribution Analysis

Feature contribution analysis using SHAP (SHapley Additive Explanations) showed that CD4+ count was the most important feature in HIV prediction, followed by CD4/CD8 ratio, age, and lymphocyte percentage. Demographic features such as gender and risk factors also contributed significantly to model accuracy. Comparison with Baseline Model. Comparisons were performed with several traditional machine learning algorithms:

Table 2. Model Performance Comparison

Model	Accuracy	Sensitivity	Specificity	F1-score
Deep Learning	94.2%	92.8%	95.1%	93.5%
Random Forest	89.7%	87.3%	91.2%	88.9%
SVM	88.4%	85.9%	90.1%	87.8%
Logistic Regression	85.2%	82.7%	87.1%	84.6%

A comparison of a model using only laboratory data with a model integrating demographic data showed a 3.7% performance improvement in accuracy and a 4.2% improvement in sensitivity. This confirms the hypothesis that integrating demographic data can improve the accuracy of HIV prediction.

The findings demonstrate that a deep learning algorithm that incorporates demographic and laboratory data can test for HIV with high accuracy. Deep learning's capacity to handle multidimensional data and recognize intricate patterns is validated by the model's better performance when compared to conventional machine learning techniques.

As the most crucial characteristic, the CD4+ count supports medical research that shows a fall in the count is a major sign of HIV infection. Nonetheless, the significance of a comprehensive strategy to HIV screening is illustrated by the contribution of demographic characteristics including age and risk factors.

The model's 92.8% sensitivity shows that it can identify HIV-positive cases, which is important because false negatives might affect the transmission of infection. In order to minimize false positives, which might worry patients and waste resources, high specificity (95.1%) is also essential.

3.6. Model Robustness Analysis to Data Variation

We carried out further tests to evaluate the model's resilience under diverse data settings that mirror actual clinical scenarios, in addition to the regular performance evaluations.

3.6.1. Model Performance with Missing Data

We performed simulated tests adding different percentages of missing data (10%, 20%, and 30%) to the test dataset in order to assess the model's resilience to missing data. According to the findings, the model's performance remained comparatively constant up to 20% missing data (91.3% accuracy, a 2.9% drop from baseline). Accuracy dropped to 87.8% at 30% missing data, yet it is still superior than conventional machine learning models with full data.

The CD4+ count and CD4/CD8 ratio were the laboratory parameters most susceptible to missing data; in these cases, missing data reduced accuracy by as much as 5.7%. In contrast, demographic data showed higher robustness, with an accuracy loss of less than 2.1% even when 30% of demographic data was missing. This attests to the multimodal approach's resilience in combining the two kinds of data.

3.6.2. Robustness to Noisy Data

To assess the model's robustness to noisy data, we introduced Gaussian noise with varying standard deviations ($\sigma = 0.1$, 0.2, and 0.3) to the laboratory data. At low noise levels ($\sigma = 0.1$), the model maintained 92.5% accuracy. At medium noise levels ($\sigma = 0.2$), accuracy decreased to 89.7%, and at high noise levels ($\sigma = 0.3$), accuracy dropped to 85.3%.

Comparison with the baseline model shows that the deep learning model has better robustness to noisy data than the Random Forest model (8.9% vs. 12.4% reduction at $\sigma = 0.2$) and SVM (8.9% vs. 13.7%). The dropout architecture implemented in our model plays a significant role in improving noise robustness.

3.6.3. Data Distribution Variation Analysis

We also evaluated the model's performance across different subpopulations to assess its generalizability. The model demonstrated good performance consistency across age groups, with accuracy variation of less than 3.2%. However, sensitivity dropped to 89.5% in the older group (>60 years), indicating a small deterioration in performance.

Geographic distribution-based analysis revealed greater variation, with accuracy varying between regions between 91.3% to 95.7%. This disparity is likely attributable to changes in HIV epidemiology trends and risk factors between locations.

3.6.4. Practical Implications

The model's application in actual clinical situations will be significantly impacted by this robustness examination. The model's potential for use in healthcare facilities with constrained data collection and processing capability is demonstrated by its ability to retain high performance even with poor data.

An adaptive preprocessing methodology was created and included to the implementation pipeline in order to maximize performance under various data situations. This approach incorporates data augmentation techniques to increase resistance to noise and a dynamic imputation method that adjusts to particular patterns of missing data.

Table 3. Model Performance on Various Data Conditions

Data Condition	Accuracy	Sensitivity	Specificity	F1-score
Baseline	94.2%	92.8%	95.1%	93.5%
Missing Data 10%	93.1%	91.4%	94.2%	92.3%
Missing Data 20%	91.3%	89.7%	92.5%	90.5%
Missing Data 30%	87.8%	85.9%	89.1%	87.2%
Noise Level $\sigma = 0.1$	92.5%	90.8%	93.7%	91.9%

4. CONCLUSION

By combining laboratory and demographic data, this team was able to create a deep learning model for HIV screening that achieved 94.2% accuracy. The model can increase the effectiveness of HIV screening in medical institutions and showed better performance than conventional machine learning methods. It was demonstrated that adding demographic information improved model accuracy by 3.7%, underscoring the significance of a multifaceted strategy in medical prediction.

This study includes a number of limitations that should be taken into account, even though the results are encouraging. Only 17 referral hospitals were included in the dataset, and since the majority of these are found in cities, they might not accurately represent the population in rural areas with varying sociodemographic traits. Furthermore, the distribution of demographic features in the dataset does not entirely match the national demography, with a larger representation of the productive-age population and particular risk categories. Another disadvantage is the possibility for bias in earlier treatment data, which may alter some laboratory measurements.The practical implementation of this model also faces challenges related to information technology infrastructure, which varies across healthcare institutions in Indonesia, as well as data privacy issues that require specific security protocols in accordance with national health regulations.

The developed model can be an effective tool for healthcare professionals in conducting faster and more accurate HIV screening. For further research, it is recommended to use larger and more diverse datasets from various regions in Indonesia, integrate additional data such as medical histories and other supporting test results, develop ensemble models that combine various deep learning algorithms, conduct external validation with datasets from various hospitals, and develop mobile or web applications for practical implementation in healthcare facilities.

Implementing this model in hospital information systems could improve the effectiveness of HIV prevention and control programs in Indonesia, particularly in areas with limited laboratory access. However, careful consideration of the model's limitations and adaptation to the local conditions of each healthcare facility is essential.

REFERENCES

Beam, A. L., & Kohane, I. S. (2018). Big data and machine learning in health care. JAMA, 319(13), 1317-1318. https://doi.org/10.1001/jama.2017.18391

Branson, B. M., Owen, S. M., Wesolowski, L. G., Bennett, B., Werner, B. G., Wroblewski, K. E., & Pentella, M. A. (2014). Laboratory testing for the diagnosis of HIV infection: Updated recommendations. Centers for Disease Control and Prevention. MMWR Recommendations and Reports, 63(RR-06), 1-25.

Ching, T., Himmelstein, D. S., Beaulieu-Jones, B. K., Kalinin, A. A., Do, B. T., Way, G. P., ... & Greene, C. S. (2018). Opportunities and obstacles for deep learning in biology and medicine. Journal of The Royal Society Interface, 15(141), 20170387. https://doi.org/10.1098/rsif.2017.0387

Cohen, M. S., Chen, Y. Q., McCauley, M., Gamble, T., Hosseinipour, M. C., Kumarasamy, N., ... & Fleming, T. R. (2011). Prevention of HIV-1 infection with early antiretroviral therapy. New England Journal of Medicine, 365(6), 493-505. https://doi.org/10.1056/NEJMoa1105243

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115-118. https://doi.org/10.1038/nature21056

Kementerian Kesehatan Republik Indonesia. (2023). Laporan Perkembangan HIV-AIDS dan Penyakit Infeksi Menular Seksual (PIMS) Triwulan IV Tahun 2023. Jakarta: Direktorat Jenderal Pencegahan dan Pengendalian Penyakit.

Kumar, S., Agarwal, M., & Sharma, A. (2020). Machine learning approaches for diabetes prediction using integrated clinical and demographic data. Journal of Medical Internet Research, 22(8), e19428. https://doi.org/10.2196/19428

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444. https://doi.org/10.1038/nature14539

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. New England Journal of Medicine, 380(14), 1347-1358. https://doi.org/10.1056/NEJMra1814259

Zhang, L., Wang, Y., Niu, M., Wang, C., & Wang, Z. (2019). Machine learning for characterizing risk of type 2 diabetes mellitus in a rural Chinese population: The Henan Rural Cohort Study. Scientific Reports, 9(1), 4406. https://doi.org/10.1038/s41598-019-40928-9