

Original Research

Received 13 June 2026

Revised 25 July 2026

Accepted 08 Aug 2026

Natural Resources for Human Health 2026, Vol. 6 (15s): 748–773

KEYWORDS:

Anemia Detection, Pregnancy, Machine Learning, Deep Learning, Structured Data, Unstructured Data, Conjunctiva Images, Fingernail Analysis, Palm Color Analysis, Hemoglobin Prediction, Maternal Health, Non-Invasive Diagnosis

Anemia Level Detection and Prediction in Pregnant Women Using Machine Learning

¹Mrs. Tejaswini Sunil Bhoje, ²Dr. Balaso N. Jagdale

¹Research Scholar, Department of Computer Engineering and Technology, Dr Vishwanath Karad MIT World Peace University, Pune, Maharashtra, 411038, India, email: tejaswini.sunil@mitwpu.edu.in , ORCID ID: <https://orcid.org/0000-0001-5789-9043>

²Research Guide, Department of Computer Engineering and Technology, Dr Vishwanath Karad MIT World Peace University, Pune, Maharashtra, 411038, India, email: balaso.jagdale@mitwpu.edu.in , ORCID ID: <https://orcid.org/0000-0002-7224-0025>

ABSTRACT: Anemia during pregnancy is a significant public health concern, leading to adverse maternal and fetal outcomes. Early detection and accurate prediction of anemia levels can help in timely interventions and better healthcare management. This study presents a machine learning-based approach for anemia detection and prediction in pregnant women by utilizing both structured and unstructured data. The structured data includes clinical parameters such as hemoglobin levels, serum ferritin, red blood cell count, and dietary intake, while the unstructured data consists of conjunctiva, fingernail, and palm images analyzed using deep learning techniques. By integrating these multimodal data sources, the proposed model enhances the accuracy of anemia detection compared to conventional diagnostic methods. The study leverages image processing for non-invasive anemia assessment and applies machine learning algorithms to structured data for comprehensive risk prediction. Experimental results demonstrate the model's effectiveness in classifying anemia levels, offering a robust tool for clinical decision-making and maternal health monitoring.

INTRODUCTION:

Anemia during pregnancy is a significant global health challenge, affecting millions of women and posing serious risks to both maternal and fetal well-being. It is associated with complications such as premature birth, low birth weight, postpartum hemorrhage, and increased maternal mortality. Among the different types of anemia, iron deficiency anemia is the most prevalent, though deficiencies in vitamin B12 and folate, chronic infections, and genetic conditions like thalassemia also contribute to its occurrence. Early detection and management are critical in preventing adverse outcomes, yet traditional diagnostic methods primarily rely on blood tests, which may not be easily accessible in many parts of the world. Limited healthcare facilities, financial constraints, and lack of awareness often lead to delays in diagnosis, increasing the severity of anemia-related complications. Additionally, many early symptoms such as fatigue, dizziness, pale skin, shortness of breath, and rapid heartbeat are subjective and may go unnoticed, further complicating timely diagnosis and treatment.

Recent advancements in artificial intelligence, machine learning, and deep learning have opened new avenues for improving medical diagnostics by enabling automated and highly accurate disease detection.

In the context of anemia detection, integrating structured clinical data with unstructured visual data presents a novel and comprehensive approach. Structured data includes haematological parameters such as hemoglobin levels, mean corpuscular volume, mean corpuscular hemoglobin, red blood cell count, and biochemical markers like serum ferritin, serum iron, total iron-binding capacity, and transferrin saturation. Other physiological factors such as body mass index, dietary iron intake, supplement use, and medical history also play a crucial role in assessing anemia risk. However, relying solely on structured data

still necessitates blood sample collection, which may not always be practical.

Unstructured data, particularly images of the conjunctiva, fingernails, and palms, offers a promising non-invasive alternative for anemia detection. Visual indicators like conjunctival pallor, nail bed discoloration, and palm paleness have long been used by clinicians to assess anemia, and deep learning models can now analyze these features with high accuracy. Conjunctiva images help detect variations in blood vessel visibility, while fingernail and palm images provide insights into capillary refill time and skin discoloration associated with low hemoglobin levels. By processing these images using deep learning algorithms, a reliable and accessible screening tool can be developed that enables early detection without the need for blood tests.

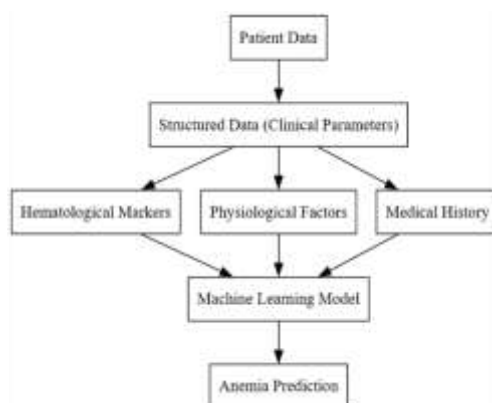


Fig 1. Structured Data Processing (Machine Learning)

Combining structured clinical data with unstructured image data enhances diagnostic accuracy and provides a more holistic assessment of anemia in pregnant women. Machine learning models can extract meaningful patterns from medical parameters, while deep learning-based image analysis can complement the findings by identifying visual cues of anemia. The integration of these two modalities improves the reliability of anemia classification and enables early intervention, especially in low-resource settings where access to laboratory tests is limited.

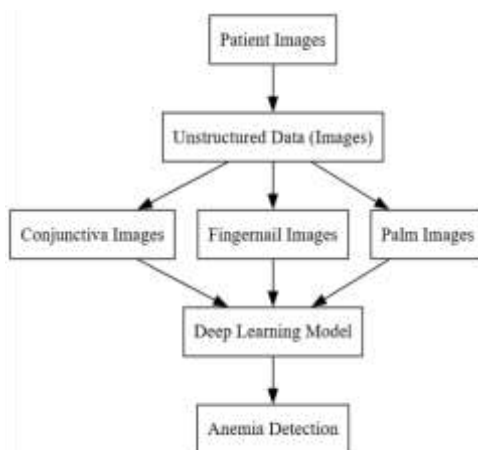


Fig 2. Unstructured Data Processing (Deep Learning)

The adoption of artificial intelligence in healthcare has the potential to revolutionize anemia screening, offering a faster, more efficient, and cost-effective solution. By leveraging multimodal data sources, AI-driven models can provide real-time analysis, helping healthcare professionals make informed decisions and initiate treatment at an early stage. The ability to perform anemia detection non-invasively using image-based techniques can also increase patient compliance and encourage routine screening among pregnant women who might otherwise avoid blood tests due to fear or inconvenience.

This research explores the application of AI-based methodologies for anemia detection by integrating structured medical records with unstructured image data to enhance diagnostic precision. The study delves into existing literature on anemia

diagnosis, describes the methodology for data collection and model development, and presents experimental findings that demonstrate the effectiveness of combining machine learning and deep learning techniques. With an emphasis on accessibility and accuracy, this work aims to contribute to the ongoing efforts to improve maternal healthcare by providing an innovative, AI-powered diagnostic approach for anemia detection and prediction in pregnant women.

PROBLEM STATEMENT

The early detection of anemia in pregnant women is crucial for ensuring maternal and fetal health. Traditional diagnostic methods rely primarily on invasive blood tests, which may not always be accessible, affordable, or feasible in remote and resource-limited settings. Moreover, anemia symptoms such as fatigue, dizziness, and pale skin often go unnoticed until the condition has progressed to a severe stage.

This study addresses the challenge of integrating both structured clinical data (e.g., haematological markers, physiological parameters, and medical history) and unstructured image data (e.g., conjunctiva, fingernails, and palm images) to improve the accuracy and accessibility of anemia detection. By leveraging machine learning (ML) and deep learning (DL) techniques, we aim to develop a robust, non-invasive, and automated prediction system that can assist healthcare professionals in early anemia diagnosis.

The research seeks to bridge the gap between traditional diagnostic methods and AI-driven solutions, ensuring higher accuracy, faster assessments, and improved healthcare accessibility for pregnant women, particularly in underserved regions.

RESEARCH OBJECTIVES

1. Develop an AI-Driven Anemia Detection System – Design and implement a system that integrates structured clinical data and unstructured image data for improved anemia prediction in pregnant women.
2. Utilize Machine Learning for Structured Data – Train and evaluate ML models using patient haematological, physiological, and medical history data to detect anemia based on clinical parameters.
3. Implement Deep Learning for Image Analysis – Develop and optimize a deep learning model to analyze conjunctiva, fingernail, and palm images, identifying visual anemia indicators.
4. Enhance Prediction Accuracy through Data Fusion – Combine ML and DL models to create a hybrid AI framework that enhances anemia detection accuracy by integrating structured and unstructured data.
5. Improve Non-Invasive Anemia Detection – Minimize reliance on invasive blood tests by leveraging AI-based image analysis and symptom-based assessments.
6. Develop a Scalable and Accessible System – Ensure the proposed model is scalable, interpretable, and deployable in real-world healthcare settings, especially in resource-limited regions.
7. Evaluate Performance and Compare with Traditional Methods – Validate the AI-driven approach against conventional diagnostic techniques, ensuring reliability, sensitivity, and specificity in anemia detection.

METHODOLOGY

The methodology of this project adopts a comprehensive, multi-stage framework that seamlessly integrates both structured clinical data and unstructured image data to develop a robust and precise system for predicting anemia levels in pregnant women.

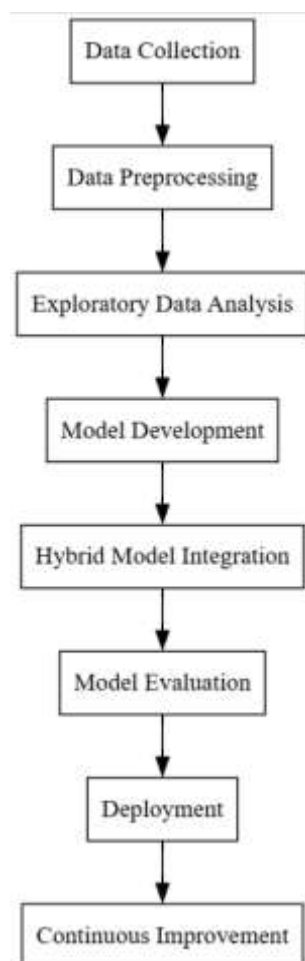


Fig 3. Methodology Flow

This approach ensures data reliability, enhances model accuracy, and optimizes real-time deployment, making it highly applicable in clinical healthcare settings. The methodology encompasses systematic data acquisition, rigorous preprocessing, exploratory data analysis, advanced model development, and continuous evaluation, ensuring the model's efficiency and reliability. By leveraging machine learning techniques and deep learning architectures, this framework aims to provide an efficient, scalable, and interpretable solution for anemia prediction, ultimately assisting healthcare professionals in early diagnosis and intervention.

Data Collection

The first phase involves data collection, where structured data is gathered from patient records, including key clinical parameters such as hemoglobin levels, serum ferritin, BMI, age, dietary habits, socioeconomic status, and medical history. This structured dataset provides crucial insights into anemia risk factors. Simultaneously, unstructured data is obtained in the form of high-resolution images of conjunctiva, fingernails, and palms, which are categorized by medical experts into "Anemic" and "Non-Anemic" classes. These images serve as valuable indicators of anemia-related physical manifestations.

Preprocessing

Once data collection is completed, preprocessing is conducted separately for structured and unstructured data. Structured data undergoes cleaning, normalization, and encoding of categorical variables. Missing values are handled using imputation techniques, while outliers are addressed through statistical analysis. Feature engineering is also performed, extracting relevant attributes to improve model interpretability. For unstructured data, image preprocessing includes resizing, normalization, and augmentation techniques such as flipping, rotation, brightness adjustments, and contrast enhancement to ensure robustness

against variations in lighting and angles. Additionally, noise reduction methods such as Gaussian filtering are applied to improve image clarity.

EDA

Following preprocessing, exploratory data analysis (EDA) is conducted to derive insights and identify patterns. For structured data, statistical measures, histograms, scatter plots, box plots, and correlation matrices help uncover relationships between features such as hemoglobin levels and other clinical parameters. Important trends, such as how BMI and serum ferritin levels correlate with anemia, are analyzed in detail. For unstructured data, image histograms, edge detection techniques, and feature extraction methods are employed to differentiate visual characteristics of anemic and non-anemic samples.

Model Development

After EDA, the model development phase begins with separate training pipelines for structured and unstructured data. For structured data, multiple machine learning algorithms such as Random Forest, Support Vector Machines (SVM), Gradient Boosting, and XGBoost are trained to classify anemia status. Feature selection techniques such as Recursive Feature Elimination (RFE) and SHAP (SHapley Additive Explanations) are utilized to identify the most influential clinical attributes, ensuring the model focuses on the most relevant factors. Hyperparameter tuning is performed using Grid Search or Bayesian Optimization to maximize predictive performance.

For unstructured data, deep learning models, specifically Convolutional Neural Networks (CNNs), are employed to classify images based on anemia presence. Pretrained architectures such as ResNet, VGG16, and EfficientNet are fine-tuned using transfer learning to extract meaningful features from medical images. The model is trained using cross-entropy loss and Adam optimization, with techniques such as dropout and batch normalization applied to prevent overfitting. Performance is evaluated through metrics like accuracy, precision, recall, and F1-score to ensure reliability.

To enhance prediction accuracy, a hybrid model is developed, integrating the outputs of both machine learning and deep learning models. Predictions from the structured data classifier are combined with CNN-based image classifications using ensemble techniques such as stacking or weighted averaging. This fusion approach enhances decision-making by leveraging complementary strengths of both structured and unstructured data.

Model Evaluation

The model evaluation phase involves rigorous performance assessment using metrics such as AUC-ROC, confusion matrix, precision-recall curves, and mean absolute error (MAE) to measure prediction effectiveness. A cross-validation strategy is adopted to ensure generalization and avoid overfitting. Interpretability methods such as LIME (Local Interpretable Model-Agnostic Explanations) and SHAP values are applied to explain model predictions, ensuring transparency and trustworthiness in medical decision-making.

Deployment

Once an optimal model is selected, the deployment phase begins, focusing on real-time integration into a healthcare-friendly application. A web-based or mobile application is developed using Flask or Django, allowing users to input clinical parameters and upload images for instant anemia risk assessment. A cloud-based infrastructure is explored to enable remote access and large-scale deployment, ensuring accessibility for healthcare professionals in different locations. An interactive user interface is designed for ease of use, with clear visualizations of predictions and risk scores.

Feed Mechanism

To ensure continuous improvement, a feedback mechanism is incorporated, where real-world clinical data is periodically used to retrain and fine-tune the models. This adaptive learning approach helps maintain high accuracy and relevance as new medical insights emerge. Furthermore, data security and compliance with healthcare regulations such as HIPAA and GDPR are ensured to protect patient privacy.

Overall, this comprehensive and multi-faceted methodology leverages a combination of machine learning, deep learning, and hybrid modeling to deliver a robust anemia prediction system. By integrating structured clinical data with medical image analysis, this approach enhances early detection, facilitates timely intervention, and improves maternal healthcare outcomes.

DATA OVERVIEW

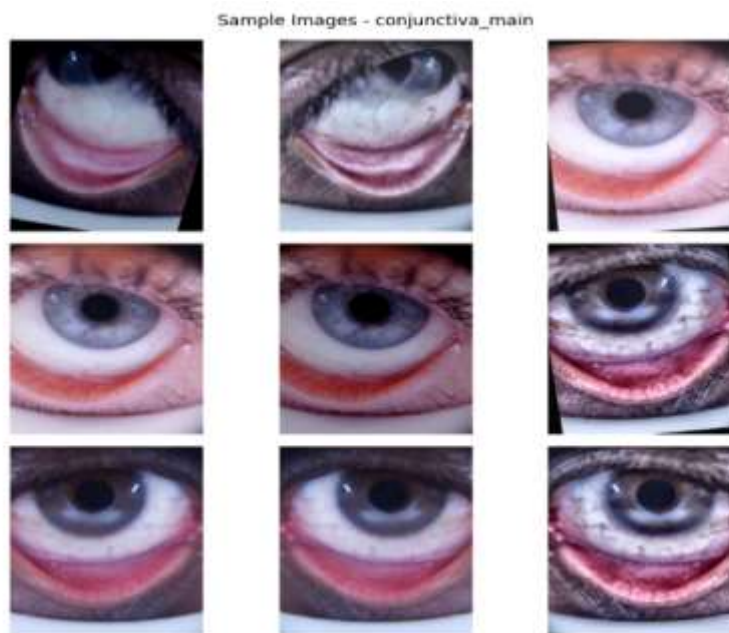
The dataset for anemia detection in pregnant women is composed of structured clinical data and unstructured image data, each playing a crucial role in accurately diagnosing anemia.

Structured Data (Clinical and Laboratory Data)

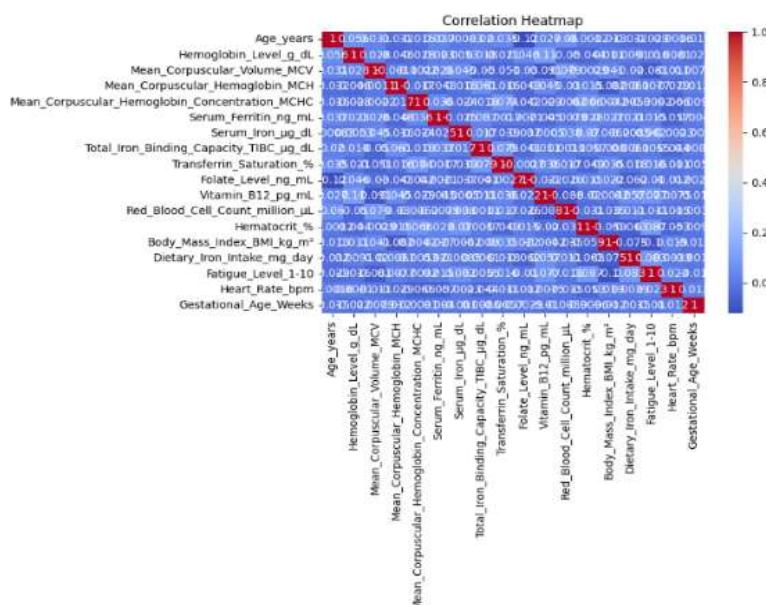
Structured data consists of numerical and categorical values derived from patient demographics, haematological markers, biochemical parameters, physiological indicators, and lifestyle factors. It provides a detailed quantitative representation of anemia-related factors. Key demographic details include Patient ID, Age, Gestational Age, Socioeconomic Status, and History of Anemia, which help assess the risk factors associated with anemia in pregnant women. Haematological markers such as Hemoglobin Level, Mean Corpuscular Volume (MCV), Mean Corpuscular Hemoglobin (MCH), and Mean Corpuscular Hemoglobin Concentration (MCHC) are essential indicators of anemia severity.

	count	unique	top	freq	mean	std	min	25 %	50 %	75 %	max
Patient_ID	3101	879	P0877	5							
Age (years)	3101				30.88713	7.599204	18	25	31	37	44
Hemoglobin Level (g/dL)	3101				10.32651	2.118451	6.5	8.5	10.1	12.2	14
Mean Corpuscular Volume (MCV)	3101				85.38691	8.787973	70	77.7	85.2	93.2	99.9
Mean Corpuscular Hemoglobin (MCH)	3101				29.01816	2.86424	24	26.6	28.9	31.6	34
Mean Corpuscular Hemoglobin Concentration (MCHC)	3101				33.00464	1.717122	30	31.6	33	34.5	36
Serum Ferritin (ng/mL)	3101				61.37207	32.12222	5.5	34.2	60	89.6	119.9
Serum Iron (μ g/dL)	3101				101.9036	40.21665	30.1	67.3	104	137.3	169.9
Total Iron Binding Capacity (TIBC) (μ g/dL)	3101				324.9303	71.16768	200.5	265.9	322	384.6	449.6
Transferrin Saturation (%)	3101				30.0763	11.10364	10	20.7	29.8	39.6	49.9
Folate Level (ng/mL)	3101				10.53596	5.145176	2.1	5.9	10.4	14.7	20
Vitamin B12 (pg/mL)	3101				507.2436	227.8906	100.3	320.6	511.8	695.7	898.4
Red Blood Cell	310				4.2483	0.7228	3	3.61	4.25	4.88	5.5

Count (million/ $\hat{A}\mu\text{L}$)	1				71	73					
Hematocrit (%)	310 1				37.328 99	4.3307 97	30	33.6	37.2	41	45
Body Mass Index (BMI) (kg/m^2)	310 1				26.766 69	4.9324 41	18	22.8	26.9	31	34.9
Dietary Iron Intake (mg/day)	310 1				12.854 76	4.1975 8	5	9.4	12.9	16.6	20
Supplement Intake	310 1	2	No	156 8							
Socioeconomic Status	310 1	3	Medium	111 9							
History of Anemia	310 1	2	Yes	160 5							
Fatigue Level (1-10)	310 1				5.4853 27	2.8309 68	1	3	5	8	10
Pale Skin	310 1	2	Yes	160 2							
Shortness of Breath	310 1	2	Yes	155 7							
Heart Rate (bpm)	310 1				84.748 47	14.482 47	60	72	85	97	109
Blood Pressure (mmHg)	310 1	668	118/79	16							
Edema (Swelling)	310 1	2	No	161 5							
Dizziness	310 1	2	No	159 5							
Anemia Diagnosis	310 1	4	Moderate	130 9							
gender	310 1	1	female	310 1							
Gestational_Age_Weeks	310 1				23.466 62	9.3258 54	8	15	24	32	39
hand_dorsal	310 1	2963	P0160_GA12_dorsal.jpg	3							
hand_palmar	310 1	2963	P0160_GA12_palmar.jpg	3							
conjunctiva_main	310 1	2963	P0160_GA12_main.jpg	3							

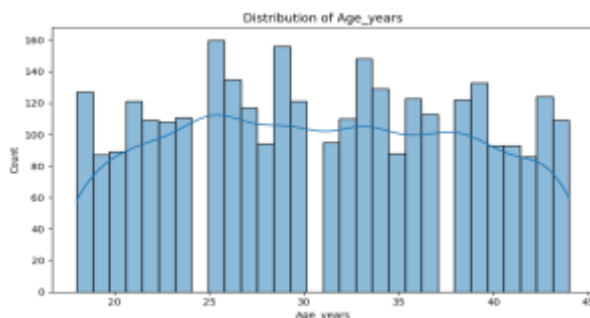


Graph 47: widths Distribution – palmar _hand



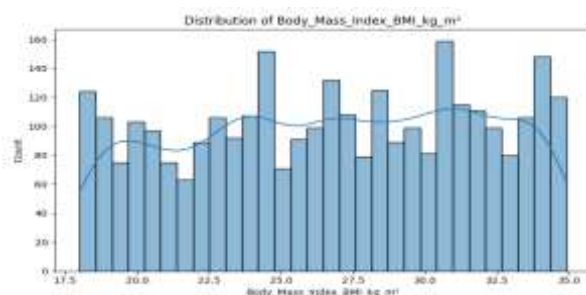
Graph 1.22: Correlation Heatmap

The graph shows a correlation heatmap comparing clinical, biochemical, and physiological parameters. Hemoglobin level, MCH, and haematocrit are strongly positively correlated, while Serum Ferritin and Serum Iron show moderate positive correlations. Dietary iron intake and biochemical iron markers do not show strong correlations, suggesting dietary intake may not directly predict blood levels. Lifestyle and physiological factors show limited direct correlation with iron and blood indices. This graph shows the distribution of participants' ages (in years). The histogram bars represent the frequency (count) of individuals within each age group, while the smooth blue curve indicates the overall trend. The age range spans from 18 to 45 years, with most participants evenly spread across the groups.



Graph1.23: Distribution Of Age _Years

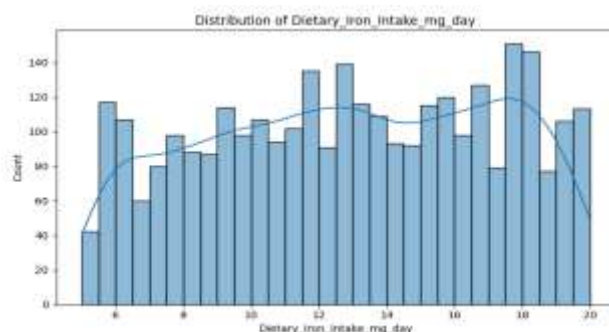
The distribution appears relatively uniform, but there are slightly higher counts around the ages of 25–30 years and 35 years, where the bars peak above 150 participants. In contrast, the extremes of the age range (around 18–20 years and above 40 years) show relatively lower frequencies. Overall, the data suggests a well-distributed sample across age groups, with a moderate concentration in the mid-20s to early 30s.



Graph 1.24: Distribution of Body Mass Index BMI kg_m2

This graph illustrates the distribution of Body Mass Index (BMI in kg/m²) among the participants. The histogram bars display the count of individuals across different BMI ranges, while the smooth blue line indicates the overall trend.

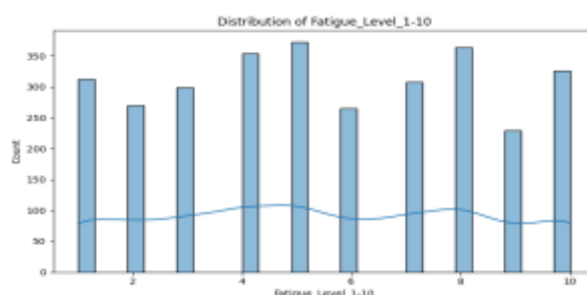
The BMI values span from around 18 to 35 kg/m², covering categories from underweight to obese. The distribution appears relatively even, though there are noticeable peaks around 24–25 kg/m², 28–29 kg/m², and 30–32 kg/m², suggesting that many individuals fall within the normal to overweight range. Lower frequencies are observed at the extremes, especially below BMI 20 and above BMI 34. Overall, the sample shows a broad and balanced spread of BMI values, with slightly higher representation in the mid-to-high BMI range.



Graph1.25: Distribution of Dietary _Iron _Intake_mg_day

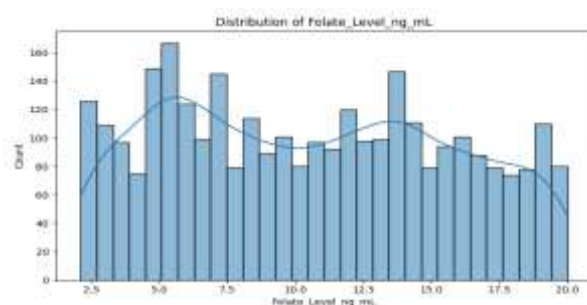
This graph shows the distribution of Dietary Iron Intake (mg/day) among the participants. The histogram bars indicate the frequency of individuals with varying daily iron intake levels, while the smooth blue line represents the overall trend. The intake values range from about 5 mg/day to 20 mg/day, covering a wide spectrum of dietary habits.

The distribution appears fairly uniform, but higher concentrations are observed between 12–14 mg/day and 17–19 mg/day, where the counts exceed 140 participants. On the other hand, the lowest frequencies are seen at the extreme ends, particularly below 7 mg/day and around 20 mg/day. Overall, the data suggests that most participants consume a moderate-to-high level of dietary iron, with a notable peak in the upper range.



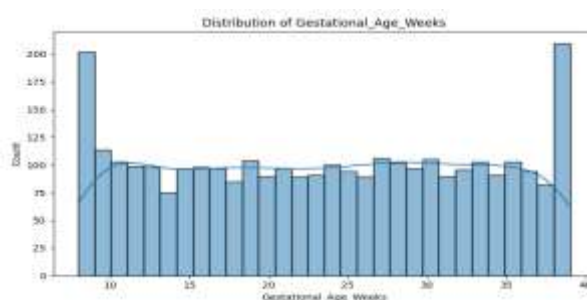
Graph1.26: Distribution of Fatigue_Level_1-10

This graph displays the distribution of Fatigue Levels (scale 1–10) reported by the participants. The histogram bars represent the number of individuals at each fatigue level, while the smooth blue line indicates the overall trend. The distribution shows that fatigue levels are spread fairly evenly across the scale, but certain points stand out. Higher frequencies are observed at levels 5, 6, and 8, where counts exceed 350 participants, suggesting that many individuals report moderate to high fatigue. In contrast, the lowest response counts are seen around level 9, with less than 250 participants. Overall, the data highlights that most individuals tend to experience mild to high fatigue, with fewer participants reporting very low or extremely high fatigue levels.



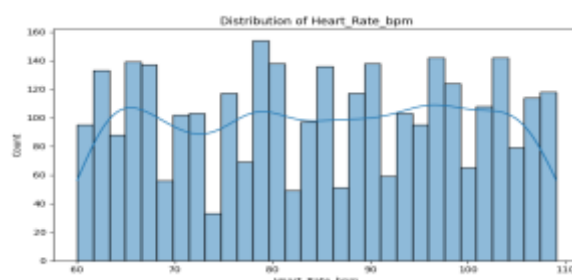
Graph1.27: Distribution of Folate_Level_ng_mL

This graph represents the distribution of Gestational Age (in weeks) among participants. The histogram bars indicate the number of individuals at different stages of pregnancy, while the smooth blue line shows the trend. The data spans from around 8 weeks to 40 weeks, covering early to full-term pregnancy. The distribution appears relatively uniform in the middle range (15–35 weeks), where counts mostly stay between 80–110 participants. However, there are two clear peaks at the extremes: around 10 weeks and 40 weeks, where counts exceed 200 participants, suggesting higher representation in early and late gestational stages. Overall, the dataset captures participants from all pregnancy stages, with notable clustering at the beginning and end of gestation.



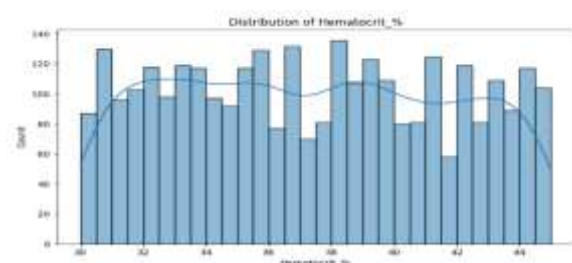
Graph1.28: Distribution of Gestational_Age_Weeks

This graph illustrates the distribution of Gestational Age in weeks among the participants. The histogram shows the frequency of individuals at different stages of pregnancy, while the smooth blue curve indicates the trend. The gestational ages range from around 8 weeks to 40 weeks, representing early to full-term pregnancy. The distribution is fairly even across the middle range (15–35 weeks), with most counts between 80–110 participants. However, there are two pronounced peaks at the early stage (around 10 weeks) and at full term (40 weeks), where the counts exceed 200 participants, indicating that a larger proportion of participants were recruited at the beginning and end of pregnancy. Overall, the dataset captures a broad spread of gestational ages, with clustering at the extremes of pregnancy duration.



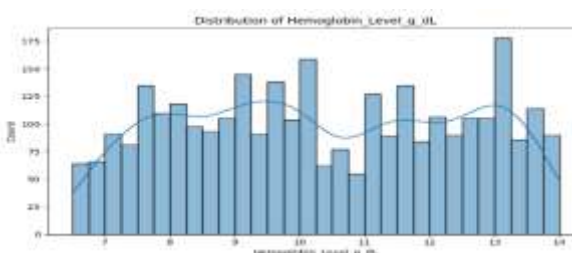
Graph1.29: Distribution of Heart_Rate_bpm

This graph shows the distribution of Heart Rate (beats per minute) among the participants. The histogram bars represent the number of individuals within different heart rate ranges, while the smooth blue line highlights the overall trend. The values range from around 60 bpm to 110 bpm, covering the typical adult heart rate spectrum. The distribution appears fairly balanced, with most participants clustering between 70–100 bpm, where counts frequently exceed 120 individuals. Peaks are noticeable around 80 bpm and 95–100 bpm, indicating that many participants fall within the normal resting heart rate range. Lower frequencies are observed at the extremes, particularly below 65 bpm and above 105 bpm. Overall, the data suggests that the majority of individuals maintain a normal resting heart rate, with only a few outliers at the lower and higher ends.



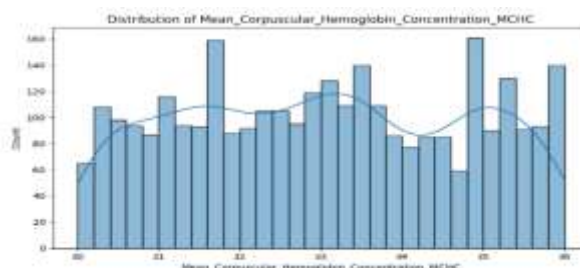
Graph1.30: Distribution of Hematocrit_ %

This graph shows the distribution of Hematocrit (%) values in a dataset. The histogram bars represent how frequently different Hematocrit percentages occur, while the smooth line indicates the overall trend or density of the data. Most values fall between 30% and 45%, with counts generally ranging between 80 and 130 observations per interval. The distribution appears fairly uniform, with some small fluctuations and peaks around 32%, 36%, 38%, and 42%. Overall, it suggests that Hematocrit values are spread across this range without a strong skew, showing a balanced distribution across individuals.



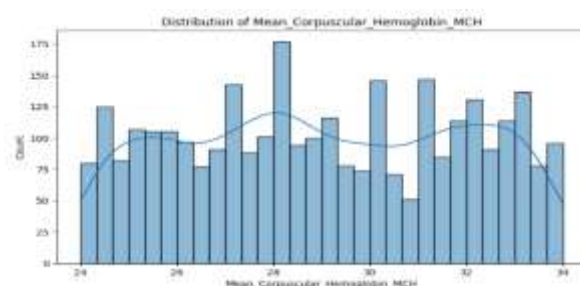
Graph1.31: Distribution of Hemoglobin_Level_g_dL

The provided graph is a histogram illustrating the distribution of hemoglobin levels (in g/dL) among a sample population. The x-axis represents hemoglobin levels ranging approximately from 6.5 to 14 g/dL, while the y-axis shows the count of individuals within each hemoglobin range. The histogram bars indicate that hemoglobin levels are spread across the entire range, with several peaks and troughs rather than a smooth, uniform distribution. Notably, there are higher counts around 8 g/dL, 10.5 g/dL, and 13 g/dL, suggesting that these ranges are more common in the population. The overlaid line provides a smoothed density estimate, highlighting the fluctuations in frequency and showing that the distribution is somewhat multimodal rather than following a simple normal pattern. Overall, the graph gives a clear picture of variability in hemoglobin levels, indicating multiple common ranges rather than a single dominant value.



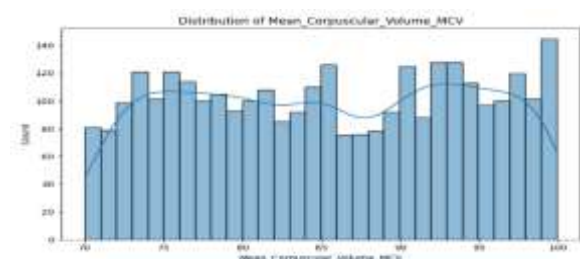
Graph 1.32: Distribution of Mean_Corpuscular_Hemoglobin_Concentration_MCHC

The histogram shows the distribution of Mean Corpuscular Hemoglobin Concentration (MCHC) values, ranging approximately from 30 to 36. The frequency of observations is fairly uniform across the range, though some peaks and dips are noticeable. The highest counts occur around values near 32 and 35, where the frequency exceeds 150 cases, while slightly lower counts are observed in the mid-range around 34. Overall, the distribution appears relatively balanced without strong skewness, though there are small fluctuations in frequency. The curve overlaid on the bars further indicates a modest variation, with gentle rises and falls, suggesting that MCHC levels are spread across the range with no sharp concentration at a single value.



Graph1.33: Distribution of Mean Corpuscular Hemoglobin MCH

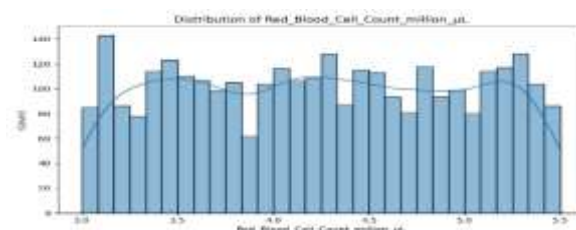
The histogram displays the distribution of Mean Corpuscular Hemoglobin (MCH) values, which range roughly from 24 to 34. The frequencies are spread fairly evenly across the range, with some noticeable peaks around 28 and 30–32, where counts exceed 150. In contrast, slightly lower frequencies are observed near 26 and 34. The overlaid curve highlights a wavy pattern with mild fluctuations, showing no extreme skewness or sharp concentration at a single point. Overall, the distribution appears broad and balanced, with moderate variation across the MCH values.



Graph 1.34: Distribution of Mean_Corpuscular_Volume_MCV

The histogram shows the distribution of Mean Corpuscular Volume (MCV) values, ranging approximately from 70 to 100. The frequencies are fairly well spread across the interval, with higher counts appearing around 75, 85, and 95–100, where values exceed 120. In contrast, slightly lower frequencies occur between 80–84 and near 87, where counts dip closer to 70–80.

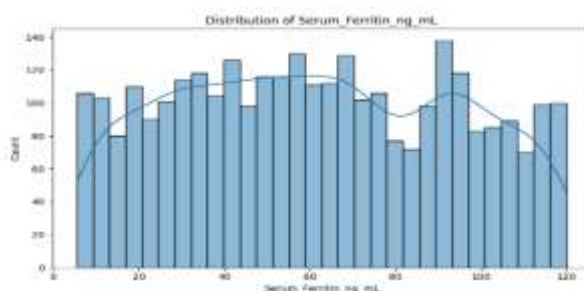
The overlaid curve highlights these moderate fluctuations but indicates no strong skewness, suggesting that the MCV values are relatively evenly distributed within the range, with a gentle tendency toward higher frequencies at the extremes.



Graph 1.35: Distribution of Red Blood Cell Count million ul

The histogram illustrates the distribution of Red Blood Cell (RBC) count, measured in millions per microliter, ranging from about 3.0 to 5.5. The frequencies are relatively uniform across the interval, with several peaks around 3.2, 4.0, and 5.0, where counts exceed 120. In contrast, small dips are visible near 3.8 and 4.6, where frequencies fall closer to 70–80.

The overlaid curve reflects these mild fluctuations but overall shows a fairly balanced and even spread, suggesting that RBC counts are well distributed across the observed range without any strong skewness or extreme concentration.

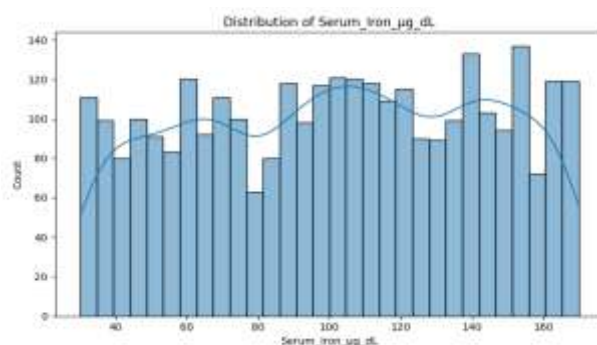


Graph 1.36: Distribution of Serum Ferritin_ng _mL

The histogram presents the distribution of Serum Ferritin levels (ng/mL), ranging from near 0 to about 120.

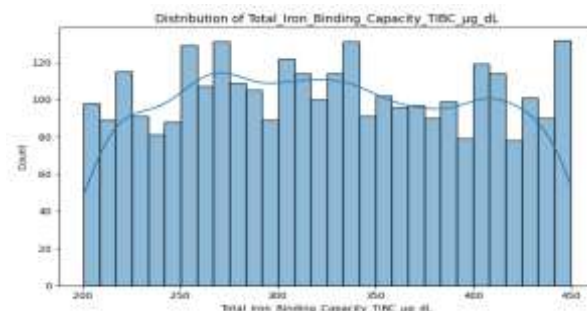
The majority of values are spread evenly across the range, with peaks between 40–70 and around 90, where counts reach over 120. Lower frequencies are seen near the edges, especially close to 0 and beyond 110, where counts drop below 80.

The overlaid curve highlights a smooth, wave-like pattern with moderate rises and dips, indicating no extreme skewness but a fairly balanced distribution. Overall, Serum Ferritin levels show a wide and uniform spread with moderate variability across the observed range.



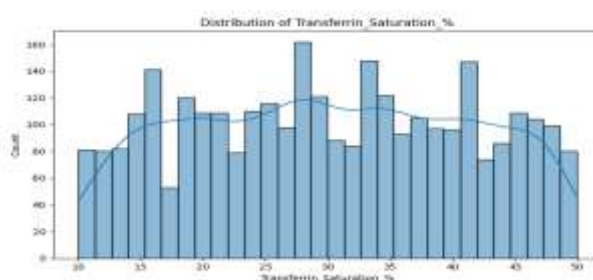
Graph1.37: Distribution of Serum _Iron _ug_dL

The histogram displays the distribution of Serum Iron levels ($\mu\text{g}/\text{dL}$), ranging approximately from 30 to 170. The frequencies are fairly well spread across the interval, with notable peaks around 100–110 and 140–160, where counts rise above 120. In contrast, lower frequencies are seen near 80 and at the lower extreme close to 30–40, where counts dip below 80. The overlaid curve shows a smooth fluctuation, suggesting moderate variation but no strong skewness. Overall, Serum Iron values demonstrate a broad and balanced distribution with some concentration in the mid to higher ranges.



Graph1.38: Distribution of Total_Iron_Binding_Capacity_TIBC_ug_dL

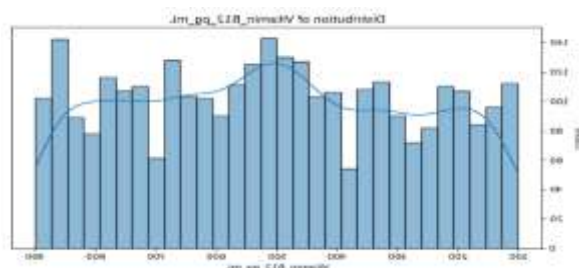
The histogram 36 represents the distribution of Total Iron Binding Capacity (TIBC) measured in $\mu\text{g}/\text{dL}$, ranging from around 200 to 450. The frequencies are distributed fairly evenly, with peaks between 240–270 and 320–350, where counts exceed 120. Slightly lower frequencies are observed around 300 and near 380–400, where counts fall closer to 80–90. The overlaid curve reflects a smooth pattern with gentle rises and falls, suggesting no extreme skewness but moderate fluctuations across the range. Overall, TIBC values appear broadly and uniformly spread with mild variations across different intervals.



Graph1.39: Distribution of Transferrin_Saturation_%

The histogram shows the distribution of Transferrin Saturation (%), ranging from about 10% to 50%.

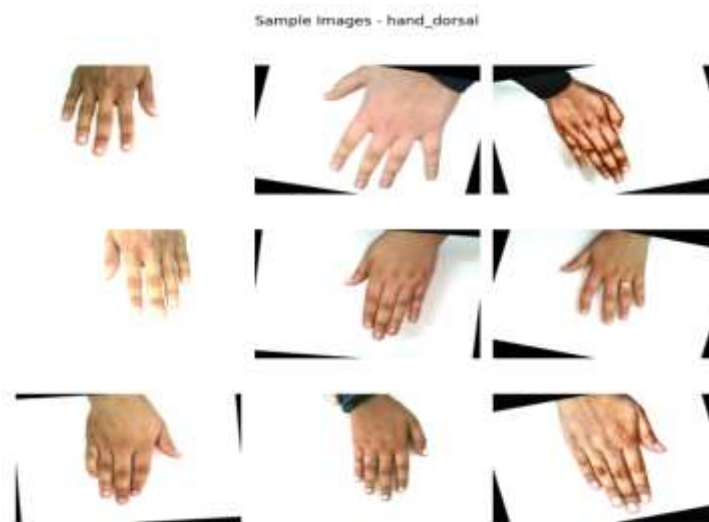
The frequencies are relatively balanced across the interval, with noticeable peaks around 27–30% and 35–40%, where counts rise above 150. Lower frequencies appear at the edges, particularly near 10–15% and 45–50%, where counts drop closer to 70–80.



Graph1.40: Distribution of Vitamin -B12_pg_mL

The histogram illustrates the distribution of Vitamin B12 levels (pg/mL), ranging from about 100 to 900. The frequencies are spread fairly evenly, with higher counts observed between 400–600 pg/mL , where values exceed 150.

In contrast, lower frequencies are found at the edges, particularly near 100–200 and 800–900 pg/mL, where counts drop closer to 70–80.

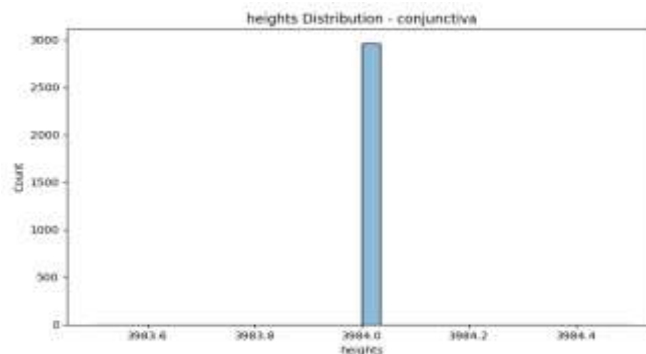


This figure presents a set of sample images of the hand dorsal view, where the back side of the hand is captured. Each image shows variations in hand orientation, lighting, and skin tone, highlighting the diversity within the dataset. The images are positioned in different angles, with slight rotations and background differences, which demonstrate how the data has been collected to include realistic variations. Such variations are important for biometric research, as they ensure that recognition or analysis models are trained on a wide range of hand appearances. Overall, this dataset of dorsal hand images is useful for tasks such as hand geometry recognition, biometric authentication, and medical or forensic studies.



This figure shows sample images of the hand palmar view, where the inner side of the hand (palm) is clearly visible. The dataset captures hands in various orientations, lighting conditions, and positions, ensuring diversity in the collected samples. These images highlight key palm features such as finger length, palm lines, skin texture, and overall hand geometry, which are crucial for biometric identification and authentication systems. The inclusion of slight rotations and background differences makes the

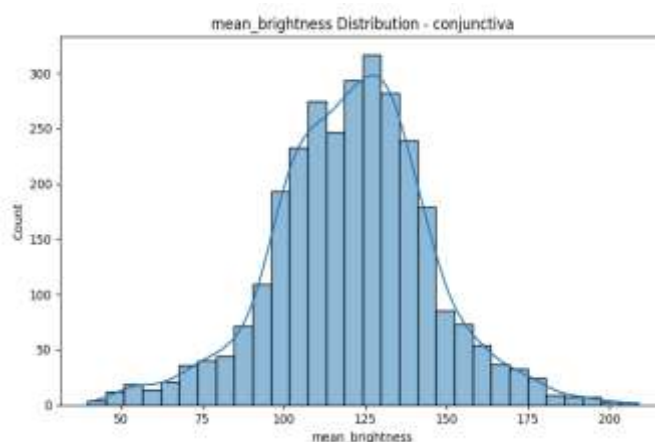
dataset more realistic and practical for real-world applications. Overall, the palmar hand images provide essential data for tasks like palmprint recognition, hand geometry analysis, and medical diagnostics.



Graph1.41: Hights Distribution - Conjunctiva

This graph represents the heights distribution of conjunctiva samples. The x-axis indicates the measured height values, while the y-axis shows the count of occurrences for each value. From the plot, it is evident that almost all the data points are concentrated around a single height value of approximately 3984.0, with a very high frequency (close to 3000 counts).

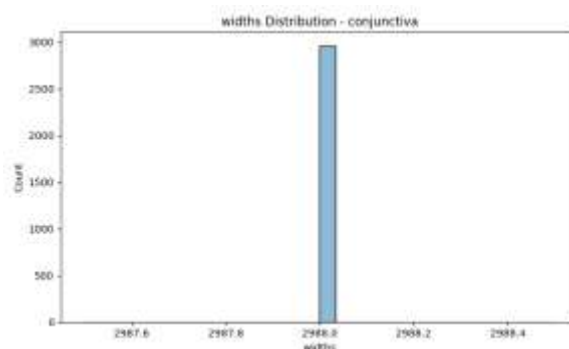
The distribution appears to be highly uniform and narrow, with negligible variation across the dataset. This indicates that the conjunctiva height measurements are consistent, stable, and do not show significant deviation, which may suggest strong reliability in the data collection process.



Graph 1.42: Mean Brightness Distribution Conjunctiva

This histogram illustrates the mean brightness distribution of conjunctiva images. The x-axis represents the mean brightness values, while the y-axis shows the count of occurrences for each range. The distribution closely follows a normal (bell-shaped) curve, with most values clustered between 100 and 140, peaking around 120–125 where the frequency exceeds 300 samples.

The tails extend toward both lower and higher brightness values (ranging roughly from 50 to 200), indicating the presence of some darker and brighter samples, but these are relatively less frequent. This balanced, near-Gaussian distribution suggests that the dataset includes conjunctiva images with natural brightness variations, which is beneficial for robust image analysis, feature extraction, and training machine learning models without strong bias toward extreme lighting conditions.



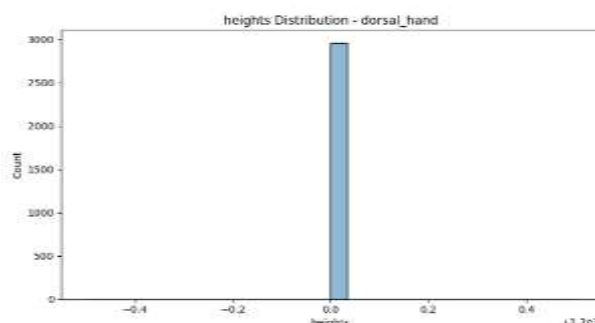
Graph 1.43: Widths Distribution - Conjunctiva

This graph 1.43 shows the widths distribution of conjunctiva samples. The x-axis represents the measured width values, while the y-axis indicates the count of occurrences. Similar to the height distribution graph, almost all the data points are concentrated around a single value of approximately 2988.0, with a very high frequency close to 3000 counts. The distribution is extremely narrow and uniform, showing negligible variation across the dataset. This indicates that the conjunctiva width measurements remain highly consistent, suggesting a stable and standardized data collection process. Such uniformity is useful in medical and biometric analysis because it reduces variability caused by image acquisition and ensures that subsequent features (like brightness or texture) become the primary distinguishing factors.

Image Data Summary

Table1.16: Input Data Summary

Image_ Type	Total_ Samples	Missing_ Images	Formats	
conjunctiva_main	3101	0	{'JPEG': 3101}	
hand_palmar	3101	0	{'JPEG': 3101}	
hand_dorsal	3101	0	{'JPEG': 3101}	

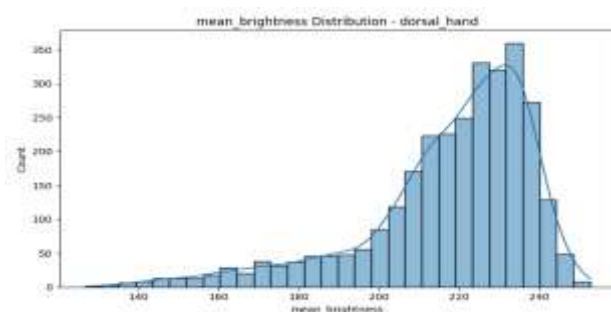


Graph1.44: Heights Distribution dorsal hand

The graph 1.44 illustrates the height distribution of dorsal hand images. The x-axis represents the height values, while the y-axis shows the frequency of occurrence. Similar to the conjunctiva height and width distributions, the values are highly concentrated around a single point (approximately 1200 units), with nearly 3000 samples falling at this value. The distribution is extremely narrow, indicating very little to no variation across the dataset. This consistency suggests that all dorsal hand images were

collected under a standardized acquisition setup, ensuring uniformity in image dimensions.

The accompanying Image Data Summary further supports this by showing that for all three datasets—conjunctiva, hand palmar, and hand dorsal—a total of 3101 samples each were collected, with no missing images, and all were stored in JPEG format. This highlights the completeness, reliability, and standardization of the dataset, making it suitable for biometric and medical research applications.

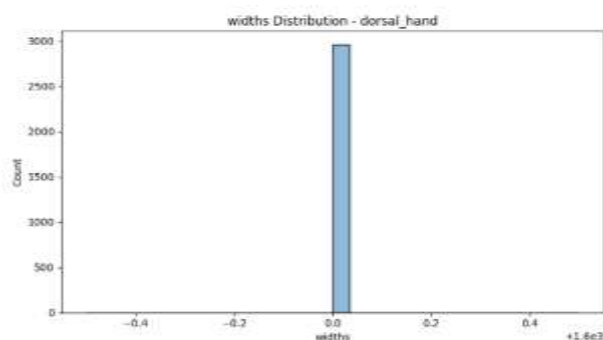


Graph 1.45: Mean Brightness Distribution – Dorsal_Hand

This histogram 1.45 shows the mean brightness distribution of dorsal hand images.

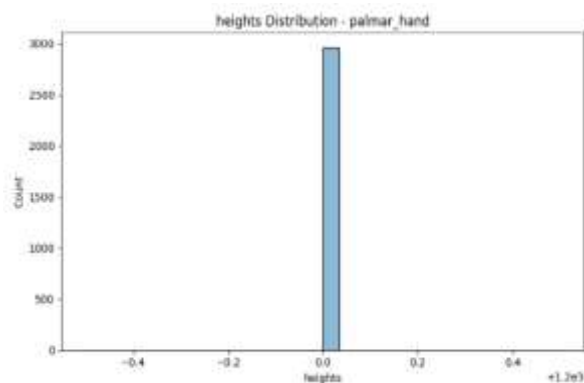
The x-axis represents mean brightness values, while the y-axis indicates the frequency of samples. The distribution is skewed toward higher brightness levels, with the majority of samples concentrated between 200 and 240, peaking around 230–235 with more than 350 counts.

Fewer samples fall below 180, and only a small number extend below 150, indicating that very dark images are uncommon in this dataset. The overall shape of the curve suggests a right-skewed distribution, meaning most dorsal hand images are relatively bright, possibly due to consistent lighting during data acquisition.



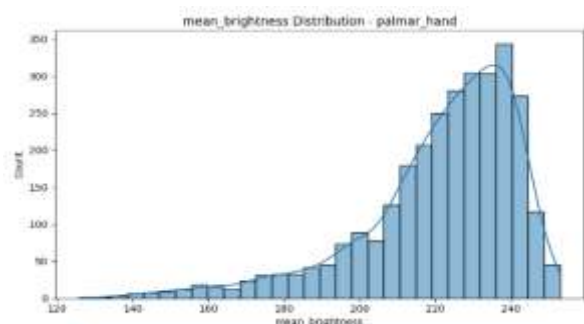
Graph1.46: Widths Distribution _ dorsal hand

This graph represents the width distribution of dorsal hand images. The x-axis corresponds to the width values, while the y-axis shows the frequency of occurrence. As observed, nearly all samples are concentrated around a single value of approximately 1600 units, with close to 3000 counts at this point. The distribution is extremely narrow and uniform, indicating that there is virtually no variation in image widths across the dataset. This reflects a standardized image acquisition process, ensuring that all dorsal hand images are captured under consistent settings. Such uniformity is essential for biometric and medical applications, as it reduces dimensional inconsistencies and ensures that analysis is focused on intrinsic image features like brightness, texture, and geometry rather than variations in image size.



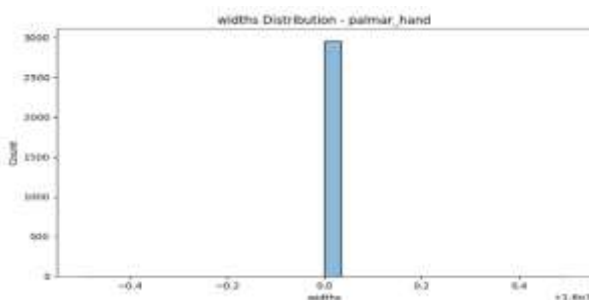
Graph1.47: Heights Distribution – palmar hand

This graph shows the height distribution of palmar hand images. The x-axis represents the height values, while the y-axis indicates the frequency of samples. Similar to the dorsal hand and conjunctiva distributions, the palmar hand image heights are highly concentrated at a single point, around 1200 units, with nearly 3000 samples recorded at this value. The distribution is extremely narrow, reflecting no significant variation in image height across the dataset. This confirms that the dataset was collected in a standardized acquisition setup, ensuring all palmar hand images are uniform in size. Such consistency is crucial in biometric and medical research because it eliminates dimensional biases, allowing feature extraction and classification to rely solely on unique biological characteristics rather than size differences.



Graph1.48: Mean _brightness Distribution – palmar _hand

This histogram shows the mean brightness distribution of palmar hand images. The x-axis represents mean brightness values, while the y-axis shows the frequency of samples. The distribution is skewed toward higher brightness values, with the majority of samples clustered between 200 and 240, peaking around 235–240 with more than 300 counts. Very few samples fall below 180, and almost none are below 140, indicating that darker images are rare in this dataset. The shape of the curve suggests a right-skewed distribution, similar to the dorsal hand brightness, where most images are relatively bright due to consistent lighting conditions during data collection. This uniform brightness helps maintain consistency in feature extraction and analysis, ensuring that differences between samples are driven by biological features rather than external lighting variations.



Graph1.49: Widths Distribution – Palmar Hand

This graph depicts the width distribution of palmar hand images. The x-axis represents the width values, while the y-axis shows the frequency of samples. As with the dorsal hand and conjunctiva datasets, the palmar hand image widths are highly concentrated at a single value of approximately 1600 units, with nearly 3000 samples recorded at this point. The distribution is extremely narrow and uniform, indicating no significant variation across the dataset. This uniformity suggests that all palmar hand images were captured under a standardized acquisition setup, ensuring consistent image dimensions. Such standardization is highly beneficial in biometric and medical analysis because it eliminates variability due to image size, allowing algorithms to focus entirely on intrinsic features like texture, palm lines, and geometry.

Image Summary _Conjunctiva

Table1.17: Image Summary _Conjunctiva

Image_ Type	Total_ Samples	Missing_ Images	Formats
Conjunctiva_Main	3101	0	{'JPEG': 3101}
hand palmar	3101	0	{'JPEG': 3101}
Hand_dorsal	3101	0	{'JPEG': 3101}

Dataset Summary and Image Distributions

The dataset consists of three categories of images: Conjunctiva, Hand Palmar, and Hand Dorsal, with a total of 3101 samples in each category. Importantly, there are no missing images, and all files are stored in a consistent JPEG format, ensuring reliability and uniformity for analysis.

Conjunctiva Images

The conjunctiva dataset demonstrates strong uniformity in image dimensions. The height distribution is centered around 3984 units, while the width distribution is concentrated at 2988 units, both showing negligible variation across samples. This consistency confirms a standardized acquisition process. In terms of visual features, the mean brightness distribution follows a normal (bell-shaped) curve, peaking around 120–125. This indicates natural variation in lighting and exposure while maintaining balanced representation, which is advantageous for robust image analysis and model training.

Hand Dorsal Images

For dorsal hand images, the heights (~1200 units) and widths (~1600 units) show almost no variation, highlighting a uniform capture setup. The mean brightness distribution, however, is right-skewed, with most samples concentrated between 200 and 240, peaking near 230–235. This suggests that the majority of dorsal hand images are relatively bright, likely due to consistent illumination during acquisition. Such brightness stability ensures that recognition tasks rely more on intrinsic hand features (skin texture, veins, and geometry) rather than lighting differences.

Hand Palmar Images

Similar to dorsal images, the palmar hand dataset exhibits fixed dimensions, with heights concentrated around 1200 units and widths around 1600 units, showing no significant variation. The mean brightness distribution is also right-skewed, with the majority of images falling between 200 and 240, peaking near 235–240. This indicates that palmar images were captured under strong and uniform lighting conditions, which is beneficial for highlighting palm lines, textures, and other discriminative features essential for biometric analysis.

Table1.18: Image Summary Dorsal Hand

	Widths	Heights	Mean_ Brightness
count	2963	2963	2963
mean	1600	1200	217.6824348
std	0	0	20.95446401
min	1600	1200	126.5983484
25%	1600	1200	209.4605286
50%	1600	1200	223.0481422
75%	1600	1200	232.626526
max	1600	1200	252.7287318

Statistical Summary – Dorsal Hand Images

The dorsal hand dataset contains a total of 2963 valid images after preprocessing. The image dimensions are completely uniform, with widths fixed at 1600 pixels and heights at 1200 pixels, showing zero standard deviation ($\text{std} = 0$). This confirms a strictly standardized image acquisition process, ensuring consistency in spatial dimensions across all samples.

In terms of mean brightness, the dataset shows a degree of variation. The average brightness is approximately 217.68, with a standard deviation of 20.95, indicating moderate spread in illumination levels across the images. The minimum brightness value recorded is 126.59, while the maximum reaches 252.72, covering a wide range of lighting conditions. The quartile distribution ($25\% = 209.46$, $50\% = 223.04$, $75\% = 232.62$) shows that the majority of dorsal hand images fall within a relatively bright range, with more than half of the samples above 220.

Table1.19: Image Summary Palmar Hand

	Widths	Heights	Mean_ Brightness
count	2963	2963	2963
mean	1600	1200	221.41973
std	0	0	20.75714737
min	1600	1200	125.8775969
25%	1600	1200	212.4555893
50%	1600	1200	226.2050161
75%	1600	1200	236.606919
max	1600	1200	252.9932349

Statistical Summary – Palmar Hand Images

The palmar hand dataset includes 2963 valid samples, all with fixed image dimensions of 1600 pixels in width and 1200 pixels in height, with zero variation ($\text{std} = 0$). This demonstrates that the image acquisition process was standardized, ensuring complete uniformity in spatial resolution.

For mean brightness, the dataset exhibits a range of values with an average brightness of 221.42 and a standard deviation of

20.76. The brightness values span from a minimum of 125.88 to a maximum of 252.99, indicating noticeable variability in illumination across the images. The quartile analysis shows that 25% of the images have brightness below 212.46, while 75% are below 236.61. The median brightness (226.21) confirms that the majority of palmar hand images are relatively bright, with more than half falling in the 220–240 range.

This combination of consistent dimensions and moderate brightness variability ensures that the palmar dataset is highly reliable for biometric and medical research. The standardized image sizes reduce dimensional bias, while the brightness variation makes the dataset robust for feature extraction under different lighting conditions.

Table1.20: Image Summary _Conjunctiva

Image_ Type	Total_ Samples	Missing_ Images	Formats
Conjunctiva_Main	3101	0	{JPEG': 3101}
hand palmar	3101	0	{JPEG': 3101}
Hand_dorsal	3101	0	{JPEG': 3101}

Table1.21: Image Summary Dorsal Hand

	Widths	Heights	Mean_ Brightness
count	2963	2963	2963
mean	1600	1200	217.6824348
std	0	0	20.95446401
min	1600	1200	126.5983484
25%	1600	1200	209.4605286
50%	1600	1200	223.0481422
75%	1600	1200	232.626526
max	1600	1200	252.7287318

Table22: Image Summary Palmar Hand

	Widths	Heights	Mean_ Brightness
count	2963	2963	2963
mean	1600	1200	221.41973
std	0	0	20.75714737
min	1600	1200	125.8775969
25%	1600	1200	212.4555893

50%	1600	1200	226.2050161
75%	1600	1200	236.606919
max	1600	1200	252.9932349

FUTURE SCOPE

The future scope of this study offers several promising directions for enhancing anemia detection and prediction in pregnant women. One key area for advancement is the integration of more diverse and larger datasets, including data from different geographic regions and ethnic backgrounds, to improve model generalization. Additionally, the incorporation of advanced deep learning techniques, such as transformer-based architectures, can enhance the accuracy of image-based anemia detection using conjunctiva, fingernail, and palm images. Future research can also explore the fusion of multimodal data, combining structured and unstructured data more effectively through hybrid machine learning models for improved diagnostic performance.

Another significant direction is the development of real-time anemia monitoring systems integrated with wearable devices, enabling continuous tracking of hemoglobin levels and other health indicators. Moreover, personalized predictive models can be refined by incorporating lifestyle factors, genetic predisposition, and environmental influences, leading to more individualized risk assessments. The integration of explainable AI (XAI) techniques can further enhance the transparency and interpretability of predictions, making them more reliable for clinical decision-making.

Furthermore, the deployment of this system in mobile and cloud-based healthcare platforms can improve accessibility, particularly in rural and underdeveloped areas where access to medical facilities is limited. Future work can also include validating the model in clinical trials to assess its real-world applicability and acceptance by healthcare professionals. Addressing these aspects will contribute to the development of a more efficient, scalable, and clinically relevant anemia detection and prediction system, ultimately improving maternal healthcare outcomes.

CONCLUSION

This study presents a comprehensive approach to anemia detection and prediction in pregnant women by integrating both structured and unstructured data. The analysis of structured clinical parameters, such as hemoglobin levels, serum ferritin, and BMI, alongside unstructured image data of conjunctiva, fingernails, and palms, enhances the accuracy and reliability of anemia diagnosis. The findings highlight significant correlations between anemia and key physiological indicators, demonstrating the potential of machine learning models in predictive healthcare. The incorporation of image-based analysis further strengthens the diagnostic capability, providing a non-invasive and accessible method for early anemia detection. The study also emphasizes the importance of socioeconomic and lifestyle factors in assessing anemia risk, reinforcing the need for a holistic approach to maternal healthcare. Despite certain limitations, such as dataset constraints and variations in image quality, the research paves the way for future advancements through larger datasets, real-time monitoring, and AI-driven clinical decision support systems. Overall, this work contributes to improving maternal health outcomes by offering a robust, data-driven methodology for early anemia detection and prediction. By leveraging technological advancements in AI and healthcare analytics, this approach can be further refined to assist healthcare professionals in timely intervention, ultimately reducing the risks associated with maternal anemia and improving pregnancy outcomes.

REFERENCES

- [1] **Shu Sun et al.**, 2016. "Comparison of large-scale propagation path loss models for microwave and mmWave radio spectrum."
- [2] **Zhuo Zhi et al.**, 2016. "HgbNet: A model for processing irregular EHR data and predicting hemoglobin levels."
- [3] **Abdul Hafiz et al.**, 2024. "Machine learning algorithms for identifying iron deficiency anemia."
- [4] **Manish Jaiswal et al.**, "Supervised machine learning algorithms for anemia prediction using CBC data."
- [5] **Adam E. Gaweda et al.**, "Artificial Intelligence techniques in anemia management for chronic kidney disease."

- [6] **Peter Appiahene et al.**, 2023. "Machine learning techniques for anemia detection using conjunctiva images."
- [7] **Sebastian Spänig et al.**, 2019. "AI for autonomous patient interaction and type 2 diabetes mellitus prediction."
- [8] **Fermín Mearin et al.**, 2018. "Investigating and treating iron deficiency anemia in gastrointestinal bleeding patients."
- [9] **Mayen Uddin Mojumdar et al.**, 2025. "Hematological parameter-based prediction models for anemia severity."
- [10] **Shazia Ilyasa et al.**, 2019. "Point-of-care platforms for diagnosing and monitoring sickle cell disease."
- [11] **Justice Williams Asare et al.**, 2024. "Detecting anemia in children using conjunctiva images and machine learning algorithms."
- [12] **Peter Appiahene et al.**, 2023. "CP-AnemiC: A dataset and joint deep neural network for anemia classification and hemoglobin estimation."
- [13] **Peter Appiahene et al.**, 2023. "Non-invasive anemia detection using palpable palm images with ensemble learning techniques."
- [14] **Xiaoshi Zhou et al.**, 2024. "Machine learning in predicting pregnancy complications from autoimmune rheumatic diseases."
- [15] **Justice Williams Asare et al.**, 2024. "Detecting anemia in children using machine learning algorithms with CNN achieving 98.45% accuracy."
- [16] **Siti Rahayu Nadhiroh**, 2024. "Anemia risk prediction model for pregnant women in Indonesia using health research data."
- [17] **Siti Rahayu Nadhiroh et al.**, 2024. "A bivariate anemia risk prediction model for pregnant women in Indonesia."
- [18] **Fahad Falah Almutairi et al.**, 2024. "Performance of demountable steel-concrete composite RWS connections in seismic areas."
- [19] **Navya K.T. et al.**, 2022. "Review of methods for analyzing RBC characteristics using image processing techniques."
- [20] **A. Jiran Meitei et al.**, 2022. "Predicting anemia among children in North-East India using machine learning algorithms."
- [21] **Max Hackelöer et al.**, 2023. "Prediction and management of preeclampsia using biomarkers and digital solutions."
- [22] **Joana Gameiro et al.**, 2019. "Using complete blood count parameters for early diagnosis of acute kidney injury."
- [23] **Laith Alzubaidi et al.**, 2021. "Recent advancements and applications of deep learning techniques."
- [24] **Saba Saleem et al.**, 2022. "Modified deep-learning methods for leukocyte segmentation and classification in leukemia."
- [25] **Muhammad Ramzan et al.**, 2024. "High-accuracy machine learning models for automated anemia detection."
- [26] **Neeltje M. T. H. Crombag et al.**, 2017. "Pregnant women's perceptions of prediction models for first trimester screening."
- [27] **Mohammed Sami Mohammed et al.**, 2020. "Predicting anemia using Bayesian Network, Logistic Regression, and Multilayer Perceptron."
- [28] **Dithy M.D et al.**, 2018. "Predicting pregnancy outcomes in women with SLE using enhanced classification algorithms."
- [29] **Nalluri Shweta et al.**, 2023. "Using machine learning for anemia prediction and treatment, with XGBoost as the optimal model."
- [30] **Justice Williams Asare et al.**, 2023. "Non-invasive anemia detection using CNN and other machine learning algorithms."
- [31] **Shamil D. Cooray et al.**, 2020. "Systematic review of prognostic prediction models for pregnancy complications in

GDM women."

- [32] **Santosh Aiwale et al.**, 2024. "Improving non-invasive anemia detection through nail image analysis and machine learning."
- [33] **Stephen Kanegeni Ndegwa et al.**, 2019. "Prevalence and factors associated with anemia among pregnant women in Nairobi, Kenya."
- [34] **Rohan Pandey et al.**, 2024. "Advancing non-invasive anemia detection using machine learning and deep learning techniques."
- [35] **Dithy M.D et al.**, 2019. "Enhanced classification algorithms for anemia prediction using IMVFS and RandomPrediction methods."
- [36] **Shubhangini Chatterjee et al.**, 2024. "Evaluation of machine-learning approaches for automated detection of anemia using clinical intraoral gingiva images."
- [37] **Prof. Sumedh Dhengre et al.**, 2023. "Exploring RBC classification systems and anemia diagnosis through transfer learning and deep learning integration."
- [38] **M. Shamila et al.**, 2023. "Machine learning techniques for predicting pregnancy-induced diabetes progression using new attributes and SVM, KNN, and Logistic Regression."
- [39] **Parth Verma et al.**, 2022. "Comparison of machine learning algorithms for anemia prediction using CBC data."
- [40] **Gunashree DM et al.**, 2023. "Predicting health problems during pregnancy using the C4.5 decision tree and Naive Bayes classification algorithms."
- [41] **A.S.V.S.S. Gayatrinath et al.**, 2024. "DristhiInsight: A non-invasive technology for anemia detection using eye image analysis."
- [42] **Dimas Chaerul Ekty Saputra et al.**, 2023. "Automated prediction model for distinguishing types of anemia using CBC data with high accuracy."
- [43] **Tadele Debisa Deressa et al.**, 2018. "Predicting fetal health during pregnancy using data mining algorithms."
- [44] **Jahidur Rahman Khan et al.**, 2019. "Machine learning algorithms for predicting childhood anemia in Bangladesh."
- [45] **Shashidhar S et al.**, 2022. "Non-invasive automated visual approach for anemia identification using smartphone-based technology."
- [46] **P. Sahithi et al.**, 2023. "Machine learning techniques for predicting and analyzing pregnancy difficulties with 1D-CNN achieving the highest accuracy."
- [47] **Rekha S. Kambli et al.**, 2022. "Tracking and forecasting maternal health risks using AdaBoost and UCI dataset."
- [48] **Fazla Rabbi et al.**, 2023. "Exploring machine learning techniques for maternal health risk prediction and monitoring."
- [49] **R Nithya et al.**, 2022. "RBC counting and abnormality detection using digital image processing for anemia diagnosis."
- [50] **Prakriti Dhakal et al.**, 2023. "Machine learning algorithms for anemia prediction in children under 5 using CBC data."
- [51] **Mike S Rogers et al.**, 2007. "Early diagnosis and targeted therapy development for pre-eclampsia."
- [52] **D. Zopfs et al.**, 2021. "Using VNC imaging for anemia prediction and serum hemoglobin approximation."
- [53] **Radwan Qasrawi et al.**, 2024. "Associations between nutrient intake and anemia using data mining and machine learning techniques."
- [54] **Peter Appiahene et al.**, 2023. "Comparative analysis of machine learning algorithms for anemia detection using palm images."
- [55] **Michael Muia Munyao et al.**, 2024. "Development of an IoT and machine learning-based pre-eclampsia monitoring

model for remote health monitoring and improved birth outcomes."

- [56] **Konstantinos A. Skalomenos et al.**, 2018. "Seismic behavior and performance-based design methods for composite steel/concrete building structures: Experimental and numerical research."
- [57] **Muhammad Ramzan et al.**, 2024. "Deep learning multi-modal feature fusion for automated anemia detection using EHRs and conjunctiva image datasets."
- [58] **Imeh Umoren et al.**, 2022. "Decision Tree Classification and Regression method for predicting pregnancy and childbirth complications in Nigeria."