Journal of Electronics, Electromedical Engineering, and Medical Informatics
Multidisciplinary: Rapid Review: Open Access Journal **Vol. 7, No. 1, January 2025, pp: 382-396; eISSN: 2656-8632**

RESEARCH ARTICLE **OPEN ACCESS**

Manuscript received August 27, 2024; revised September 22, 2024; accepted October 26, 2024; date of publication November 20, 2024
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i1.576

How to cite: Banan K. Abdulkader, Mazin H. Aziz, "Neonatal Jaundice Severity Detection from Skin Images using Deep Transfer Learning Techniques", Journal of Electronics, Electromedical Engineering, and Medical Informatics, vol. 7, no. 1, pp. 382-396, November 2024.

Neonatal Jaundice Severity Detection from Skin Images using Deep Transfer Learning Techniques

Banan K. Abdulkader , Mazin H. Aziz

Computer Engineer, University of Mosul,Mosul, IRAQ
Corresponding author: Banan K. Abdulkader (e-mail: bananaldabbagh99@ gmail.com).

ABSTRACT Neonates in the initial weeks postpartum frequently experience jaundice, a prevalent medical condition characterized by the yellow discoloration of the sclera and integumentary surfaces. This phenomenon transpires as a result of the elevation of bilirubin concentrations within the circulatory system. When bilirubin levels reach critical thresholds, they present a considerable risk for severe complications, including neurological impairment, which represents one of the gravest outcomes that may ensue if the condition is not addressed with due diligence. This study investigates a non-invasive method for assessing jaundice severity in full-term infants from 1 to 29 days, focusing on infants in Mosul city. A dataset of 344 images was collected using an iPhone 12 Pro Max (9MP camera) at Ibn Al-Atheer Hospital, capturing various skin tones and lighting conditions to ensure accurate analysis. Advanced computer vision techniques were used to classify jaundice severity into three and four categories based on skin images. Pre-trained deep transfer learning models, namely VGG16 and ResNet50, were utilized for training, with the fully connected layer removed and a suitable classifier designed for each model. VGG16 achieved 91.71% accuracy for the three-category classification, while ResNet50 reached 95.98%. For the four-category classification, accuracies of 94.92% and 94.66% were achieved, respectively. These high accuracy levels suggest that non-invasive, image-based assessments can reduce the need for repeated blood tests. This research highlights the potential of using smartphone-based methods for jaundice screening in neonatal care, providing a reliable, accessible tool to reduce strain on medical facilities and improve early detection.

INDEX TERMS Neonatal jaundice, Machine Learning, bilirubin, Total Serum Bilirubin, deep-learning, VGG16, ResNet50.

I. INTRODUCTION

Newborn jaundice is common due to increased levels of bilirubin in the blood, causing yellowing of the skin and eyes. Jaundice is one of the most prevalent medical conditions, affecting 80% of preterm neonates and between 65% to 75% of term infants. Globally, about 50% of newborns are affected by neonatal hyperbilirubinemia, with an estimated 1.1 million cases of acute jaundice occurring annually. Regions such as South Asia and Sub-Saharan Africa show the highest prevalence, where the incidence is 100 times greater in places like Nigeria compared to developed countries.

The incidence of neonatal jaundice per 1,000 live births varies widely by region, with rates between 667.8 and 738.5 in Africa, 251.3 to 473.2 in South Asia, and 3.7 to 4.4 in Europe and the Americas. This condition's incidence decreases with improvements in socioeconomic status, showing wide regional variations, with Africa having the highest rates. Neonatal jaundice remains a major public health concern, particularly in low-income areas where access to early detection tools is limited. [1], [2]. The main cause of newborn jaundice is the underdeveloped liver, which can't process excess bilirubin effectively [3]. Thus, effective methods for detecting jaundice are crucial. The assessment of bilirubin levels in the bloodstream is commonly conducted through the extraction of blood from the neonate, as shown in FIGURE 1, a process that may induce discomfort and distress. Alternatively, several non-invasive methodologies exist, such as the transcutaneous bilirubin (TcB) device, which is a non-invasive method that estimates bilirubin levels by transmitting

light through the skin, which shows strong correlations with total serum bilirubin (TSB) levels, making it a reliable

FIGURE 1. Blood Draw Procedure for Measuring Bilirubin Levels in an Infant.

screening tool[4],[5]. Also, screening for jaundice through smartphone applications shows a significant correlation with total blood bilirubin (TSB) levels, making it effective in screening for jaundice in newborns [6].

There is also skin reflectance spectroscopy. This method, integrated into smartphone applications, allows real-time color analysis, facilitating the quantification of bilirubin levels through skin reflectance [7]. The algorithms pertinent to machine learning, particularly Convolutional Neural Networks (CNN), have been widely utilized in the fields of image processing and classification [8], [9], [10], with significant emphasis on applications in the healthcare sector [10], [11], [12]. Deep learning architectures such as VGG and ResNet have experienced considerable progression over time, transforming from initial models with constrained capabilities into highly advanced structures adept at managing intricate tasks. The advancement of these models was propelled by the necessity for enhanced computational efficiency, the capacity to analyze extensive datasets, and advancements in feature extraction techniques. In the formative stages, traditional CNNs faced significant challenges, including the vanishing gradient problem and limitations in deep architectures, which catalyzed the development of more advanced structures such as VGG and ResNet.

VGG16 brought simplicity and depth to CNN architectures, using smaller filters but with more layers, making it one of the most efficient models for image classification. Conversely, ResNet50 profoundly revolutionized the domain of deep learning by integrating residual connections, which adeptly alleviated the vanishing gradient phenomenon and enabled the training of considerably deeper neural networks without detracting from performance.

These developments allowed models to manage extensive volumes of image data while preserving elevated levels of accuracy in classification endeavors. Moreover, sophisticated deep learning architectures like ResNet and VGG have assumed a crucial role in applications such as brain tumour identification [13], bone anomaly classification [14], medical image classification [15], classifying collateral circulation in stroke patients [16], and organ segmentation, with the implementation of transfer learning further enhancing their effectiveness by addressing issues related to data scarcity and overfitting [17].

These architectural frameworks, particularly VGG16 and ResNet50, were selected for the current investigation due to their remarkable effectiveness in the classification of medical imaging datasets. ResNet50's ability to mitigate vanishing gradient issues through residual connections has consistently shown excellent results in image recognition and analysis. Furthermore, the implementation of transfer learning has significantly augmented the effectiveness of these models by allowing them to achieve satisfactory performance even when utilizing limited datasets frequent challenge encountered in the domain of medical imaging. The ability of these models to learn from weights and adapt to specific medical challenges makes them particularly suitable for non-invasive techniques such as jaundice detection.

The ongoing advancement of these technologies highlights their capacity to transform clinical decision-making processes by providing accurate, non-invasive diagnostic tools. However, challenges related to interpretability and computational requirements persist, and these remain critical areas for further scholarly investigation. Several algorithms have been employed to quantify bilirubin concentrations in the bloodstream through non-invasive methodologies. One notable study, conducted by Mansour et al., utilized a specific color detection algorithm to monitor jaundice in infants. In this study, multiple images of jaundiced infants were captured from different angles, distances, and lighting conditions to improve diagnostic accuracy [18].

Olusanya et al. [19] focused on detecting newborn jaundice utilizing a color chart. They applied image processing techniques such as segmentation, pixel similarity, and white balancing to extract key pixel information using RGB (red-green-blue) values [20]. Following this, the feature extraction stage involved color mapping conversions and feature calculations to compare color change values with a specially designed 8-color calibration card on the RGB plane. In the final bilirubin level estimation stage, K-nearest neighbors and support vector regression were employed, using the features extracted earlier. The system demonstrated an 85% success rate in relation to a control group, predicated on the correlation between the outcomes derived from the proposed methodology and the results obtained from conventional blood tests.

R. Karim et al. [20] conducted a comprehensive study examining the methodologies employed for the detection and identification of jaundice inneonates through the application of machine learning and image processing techniques. The studies referenced in their analysis include focusing on the integumentary system as a physiological component. A. Gupta et al. [21], A. Althnian et al. [22], in the research [22], they analyzed data from infants with and without jaundice sourced from King Khalid University, employing methodologies such as MLP, SVM, DT, and RF. The performance metrics of recall, accuracy, and precision were documented at 64.39%, 64.77%,

and 67.39%, respectively, as derived from the amassed dataset. Conversely, when the linear, lasso, KNN, and SVR methodologies were utilized in a separate study [21], contrasting results pertaining to bilirubin levels were observed. The actual bilirubin level was approximately 15, whereas the predicted level was approximately 15.8.

Regression analysis was conducted by J. A. Taylor et al. [23] and S. Swarna et al. [24]. The focal regions in [24] included the skin of the sternum and abdomen, with images procured from India and China, amounting to a total of 35 images. The findings revealed a correlation coefficient of 0.6 for the sternum region. Conversely, the investigation in [23] exclusively targeted the sternum skin, employing a dataset comprising 350 images representative of various ethnicities in the United States. The findings indicated a sensitivity of 84.6% alongside a specificity of 75%.

The investigation conducted by J. Castro-Ramos et al. [25] primarily concentrated on the dermal tissues of the arms, forehead, palms, and soles. Employing the Support Vector Machine (SVM) technique on a set of twenty images depicting Mexican infants, the assessment unveiled a sensitivity of 71.8% coupled with a specificity of 78.8%.

In relation to the skin situated on the face, arms, feet, and central torso, linear regression methodologies were employed, as delineated by W. Y. Hsu and H. C. Cheng [26]. The study achieved an accuracy rate of 92.5%, utilizing a collection of 196 photographs sourced from Firat University.

Aydin et al. [27] gathered images of 80 infants, evenly split between jaundiced and healthy, using a 12-megapixel camera on a Samsung Galaxy Alpha smartphone. Their methodology involved placing a calibration card with eight distinct colors on the infants' abdominal area. The initial phase of the framework focused on color adjustment, ensuring that incidental light, reflections, or shading did not impact the process. Essential skin areas and their data were retained, while unnecessary regions were converted to black. The colors on the calibration card were then matched with skin details across three color spaces—RGB, YCbCr, and Lab—using feature computation and colormap conversions. These results were subsequently fed into KNN and SVR classifiers, yielding accurate results with minimal processing time.

Angelico and colleagues [28] developed user-friendly software for Android smartphones that bypassed skin color and instead employed machine learning to assess the color of infant stool using a seven-level stool color chart as a reference. The application's performance was assessed on a dataset of 165 images. The software uses the smartphone's camera to capture images of the stool and classifies infants as either normal or jaundiced. achieving sensitivity = 100%, specificity = 99.0%, and accuracy = 99.4%.

In a study published in 2023 [29], researchers developed a dataset comprising 411 images of both healthy and non-healthy infants. They used MATLAB to implement a random forest algorithm aimed at detecting jaundiced babies, but the results did not surpass those of the methods they were compared against.

The aim of this research is to develop a classification method for newborns with jaundice into several categories (three and four categories) using a mobile phone camera under normal lighting conditions, without the need for additional devices or modifications to the infant's body or the incubator. The objective is to achieve results that are comparable or competitive with those obtained through other methods.

The contributions of this study include:

1. Dataset Generation: Two datasets were created, consisting of jaundiced and healthy infants, including their ages, blood bilirubin levels, gender, and other relevant information. One dataset was collected without a color calibration card, while the other was collected using a color calibration card for comparison.
2. Enhancing Quadruple Classification: Given the weakness of this classification in previous studies, our work focused on improving its accuracy. We achieved acceptable results in this regard.
3. Achieving Competitive Accuracy: Competitive accuracy was obtained compared to previous studies, without the need for a color calibration card on the infant's body.

This method offers a practical solution that can be used by anyone with a mobile camera and standard lighting conditions, making it highly accessible and effective for jaundice screening.

II. MATERAIL AND METHODS

The objective of the study is to estimate the level of bilirubin in infants, classifying it into three or four categories based on the Bhutani chart shown in FIGURE 2.

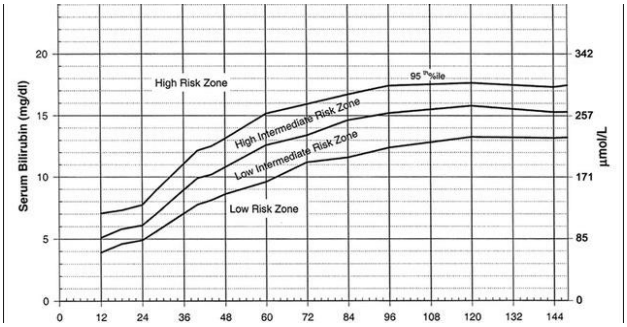

FIGURE 2. A Bhutani nomogram used to evaluate a newborn's risk by considering bilirubin levels and age, categorized into different risk zones [34].

The graph, referred to as a "Nomogram for Assessing Hyperbilirubinemia Risk," is utilized to evaluate the likelihood of jaundice in newborns by analyzing their serum bilirubin concentrations relative to their age in hours after birth, ranging from 12 to 144 hours (up to 6 days). The graphical representation is segmented into four distinct zones: the LRZ "Low Risk Zone" refers to the absence of jaundice and the minimal likelihood of its occurrence, the LIRZ "Low Intermediate Risk Zone" indicates the existence of mild jaundice, which may be managed through phototherapy, accompanied by periodic bilirubin assessments to ascertain the efficacy of the intervention, the HIRZ "High Intermediate Risk Zone" denotes a more pronounced manifestation of jaundice, necessitating rigorous phototherapy along with frequent reassessment and the HRZ "High Risk Zone" embodies a severe and critically acute instance of jaundice, which requires

extensive phototherapy and, in numerous circumstances, an exchange transfusion to mitigate bilirubin concentrations and avert complications, including neurological damage and other grave medical conditions. In this specific zone, the concentrations of bilirubin are examined with increased regularity compared to other zones.

The 95th percentile curve indicates that 95% of neonates exhibit bilirubin concentrations beneath this specific threshold, while those exceeding it are at an elevated risk for the onset of jaundice. This instrument assists healthcare practitioners in ascertaining whether a neonate necessitates additional surveillance or intervention, such as phototherapy, contingent upon their bilirubin levels in relation to their postnatal age.

A. DATASET

Due to the varying nature and color of skin among different nationalities, this study was directed toward the residents of Mosul. The database used for this research was created at Ibn Al-Atheer Teaching Hospital, with parental consent obtained for all participants. The dataset consists of images covering 386 cases of healthy and jaundiced newborns. Infants were photographed while lying on their backs inside the capsule at the neonatal intensive care unit (NICU) using a smartphone, specifically an iPhone 12 Pro Max with a 9 MP camera and an external ring light to ensure consistent lighting. The resolution of each image was 2268×4032 pixels, and the distance between the camera and the infant was approximately 30 cm, as shown in FIGURE 3.

Figure 3.Image acquisition using the mobile device camera and a supplementary lighting for an infant while he is inside the treatment incubator at the NICU.

Multiple images were captured for each infant to ensure at least one clear and usable photo. Noisy or unclear images resulting from infant movement or photographer errors were discarded, and duplicate images were also removed. Photos were taken within 5 minutes before or after blood samples were drawn from the infant. The collected images were stored in individual folders, each corresponding to a specific infant, and labeled with the sample number. Additionally, an Excel spreadsheet was created to record essential sample-specific data, including sample number, gestational age of the mother, age of the infant, and serum bilirubin (TSB) levels.

Ground truth data was meticulously collected alongside the images, incorporating key parameters such as gestational age, infant age, weight, gender, and transcutaneous bilirubin (TSB) levels. The age of the infants ranged from 1 to 22 days, with a focus on newborns aged 2 to 9 days, and infants weighing more than 2 kg. TABLE 1 provides further details regarding the infants' demographics, including their ages, genders, and weights. While the total number of images captured was 386, the number of infants ultimately included in the study was 344, comprising both healthy newborns and those diagnosed with jaundice.

TABLE 1
The additional details gathered from the infants.

Age in day	Total sample	Gender	Range of Weight kg.
Bellow 1	15	Male:9 Female:6	2.4
1	12	Male:5 Female:7	2.8
2	20	Male: 8 Female: 12	3.2
3	46	Male:24 Female:22	3
4	71	Male:38 Fmale:33	3.1
5	54	Male:27 Female:27	3.1
Above 6	168	Male:91 Female:77	3.2

Skin samples were cropped from various regions of an infant's body in order to increase the sample size and guarantee data equilibrium. As a result, the total number of newborn skin image samples that were prepared for the triple classification experiment was 2,563 divided into three classes, among whom 760 were diagnosed with severe jaundice, 803 with moderate jaundice, and 1,000 were deemed healthy. For the quadrable experiment, there were 2,558 newborns in entirety, with 760 with severe jaundice, 500 with high jaundice, 500 with low jaundice, and 798 identified as in good health as shown in FIGURE 4.

Images of infants were classified into three categories and four categories based on the Bhutan scheme. This chart includes three curves, depending on age and bilirubin level. The condition of the newborn baby under the first curve is a healthy state, while his condition between the first and second curve is a low jaundice. Between the second and third curves lies high jaundice, while surpassing the final curve leads to severe jaundice.

It is important to note that the data set for the triple classification differs from that of the quadruple classification, as each classification has its own specific data set:

1) TRIPLE CLASSIFICATION DATASET

Initially, a triple classification was used based on the Bhutani nomogram. As previously mentioned, the Bhutani nomogram

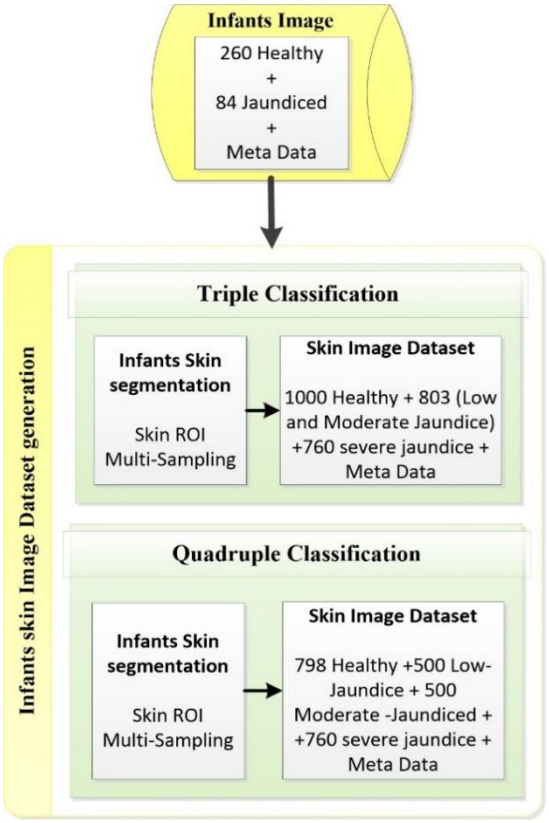

FIGURE 4.Flow diagram of the study population.

consists of three curves that create four distinct zones. In this classification, the low-intermediate risk zone and the high-intermediate risk zone were combined into a single zone named MRZ "Moderate Risk Zone," resulting in a new model comprising two curves that form three zones. The zones were merged because jaundice severity in both intermediate zones is relatively similar, with phototherapy being the standard treatment in both cases, along with regular monitoring of blood bilirubin levels. The LRZ and HRZ remained unchanged. Although the four-zone classification is more efficient and accurate, the merging was performed due to the similarities between the two intermediate zones. The regions for the triple classification were defined according to Eq. (1), Eq. (2) and Eq. (3), illustrated in FIGURE 5.

$$LRZ_{3Class} = LRZ_{4Class} \quad (1)$$

$$HRZ_{3Class} = HRZ_{4Class} \quad (2)$$

$$MRZ_{3Class} = LIRZ_{4Class} + HIRZ_{4Class} \quad (3)$$

2) QUADRUPLE CLASSIFICATION DATASET

We employed a comprehensive chart that stratifies case in to four unique areas LRZ (Low Risk Zone), LIRZ (Low Intermediate Risk Zone), HIRZ (High Intermediate Risk Zone), and HRZ (High Risk Zone) for the purposes of quadruple classification.

This chart organizes case findings to classify cases by different levels of jaundice severity, stratified by bilirubin level and other factors. However, each zone correlates to risk level, which helps in directing clinical decisions (as previously explained). Such a classification not just increases the quality of assigning severity to jaundiced infants but also guides in providing appropriate interventions based on the classified grade.

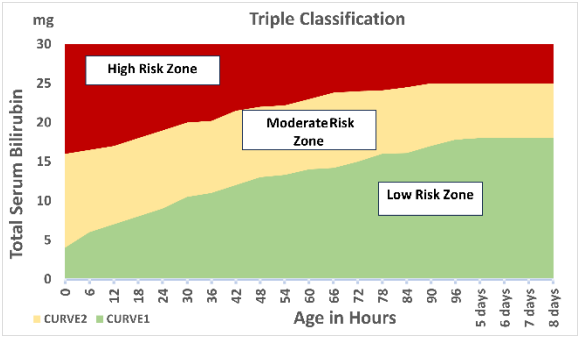

Figure 5.The Bhutani scheme is used to assess the risk of newborns by looking at bilirubin levels and age and classifying them into three different risk zones.

B. INFANT SKIN DATAEST GENERATION

Specific ROI (region of interest) zones on the child's physique were discerned. Several experiments were conducted, including only the chest area, only the abdominal area, including the forehead, and including the chest and abdomen together. It was noted that the chest and abdominal areas together gave a better result, so they were relied upon, as shown in FIGURE 6.

FIGURE 6.Skin ROI Selection examples.

These images were created using the photo editor application, and the size of the images for all children was 224 × 224. The code also resizes the images to ensure the size is 224 × 224. Several samples were taken from the same infant and given the same label as shown in FIGURE 7.

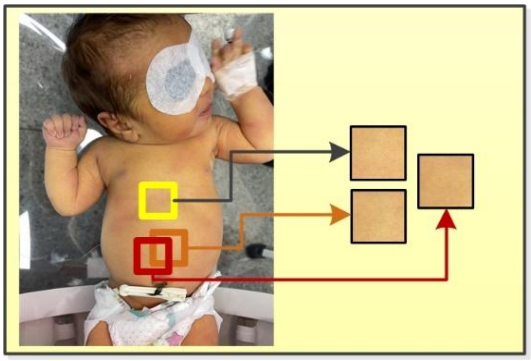

FIGURE 7.Multiple collapsed skin sampling.

C. METHODS

FIGURE 8 illustrates the process of classifying neonatal skin images into four classes: normal, low jaundice, high jaundice, and severe jaundice. A data set is established, with a division of 70% for training and 30% for testing. Deep transfer learning (DTL) using VGG16 or ResNet50 as a feature extractor with data augmentation is followed to enhance the training data further. These features are then input into a (FFNN) that is capable of performing quadratic classification. This architecture has several layers, including global mean pooling, dense layers with dropout for better generalization and the final classification of jaundice severity.

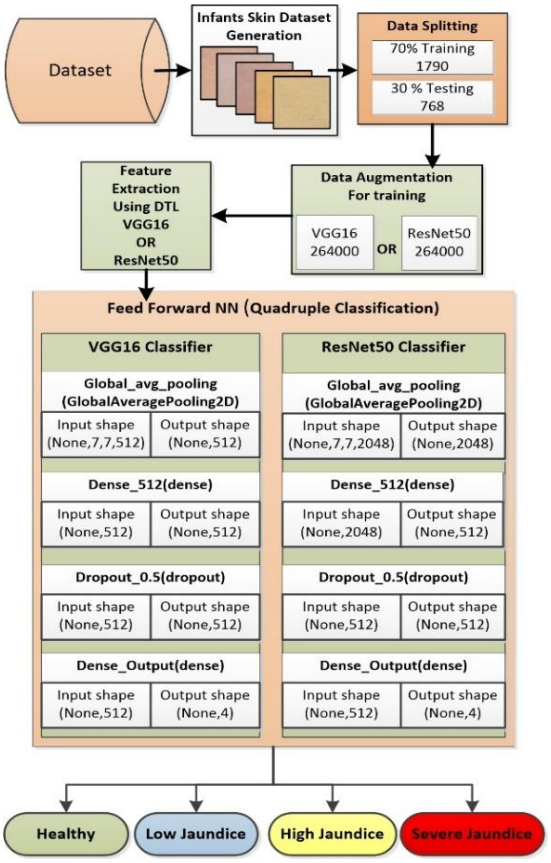

FIGURE 8.Workup block diagram for quadruple classifications.

FIGURE 9. demonstrates the process of classifying newborn skin images into three categories: Healthy, (Low and Moderate) Jaundice, and Severe Jaundice. As in the quadratic classification method, it will start by creating the dataset, then splitting it so that 70% are training images and 30% are testing images. Then the data is augmented to get the best results in training and then features are extracted using the deep learning method (DTL) we used VGG16 or ResNet50, . These features are fed to a three-class forward neural network (FFNN). It consists of layers such as normalization or global mean pooling, augmenting the dense layer with dropout to improve generalization and then classifying the images into the jaundice class.

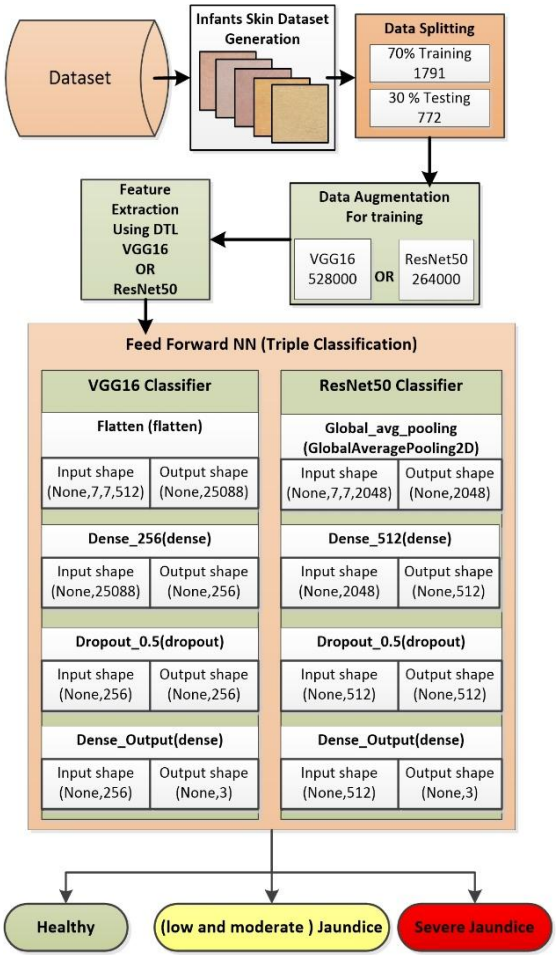

FIGURE 9.Workup block diagram for triple classifications.

1) DATA PREPARTION FOR TRAINING

The data was randomly split using Python, with 70% of the dataset allocated to the training phase and 30% to the testing phase. The next phase was then implemented using Python programming via the Google Colabe environment. Due to the scarcity of data, data augmentation was used, where seven categories of data augmentation were performed on the training data set, which included comprised rotation, width_shift, height_shift, shear, zoom, horizontal, and fill_mode.

2) FEATURE EXTRACTION

Deep transfer learning models, namely VGG16 and ResNet50, were employed to extract features from infant skin images. Each model was independently utilized for both triple and quadruple classification tasks. The pre-trained weights and parameters of the models were retained, except for the fully connected layer, where a custom classifier was designed for both triple and quadruple classifications. The models were trained with an initial learning rate of 0.001, which was dynamically adjusted during training through an optimization strategy to achieve the best learning rate. The number of epochs varied depending on the experiment, with 150 epochs used in most cases and 300 epochs for more complex training. Dropout layers were applied to enhance model generalization and reduce overfitting.

3) CLASSIFICATION

The fully connected layers were replaced by designing a SoftMax feed-forward neural network classifier having three outputs for the three categories of experiments and four outputs for the four categories of experiments for each pre-trained model (VGG16 and ResNet50). Images of infants were classified using the two algorithms into three categories and four categories. The parameters of the deep transfer learning DTL models were not changed, and the change was limited to the parameters of the classifiers. To evaluate model efficacy, metrics such as accuracy as shown in Eq. (4), precision as shown in Eq. (5), recall as shown in Eq. (6), F1-score as shown in Eq. (7), and support as shown in Eq. (8) were employed.

$$\text{Accuracy}=\frac{\text{True Positives} + \text{True Negatives}}{\text{Total Number of Instances}} \quad (4)$$

$$\text{precision}=\frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \quad (5)$$

$$\text{Recall}=\frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (6)$$

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (7)$$

$$\text{Support}=\text{Number of true instances for each class.} \quad (8)$$

III. RESULTS

A. TRIPLE CLASSIFICATION

1) VGG16

The graph of training and validation accuracy over 300 epochs demonstrates an initial rapid improvement, with the training accuracy consistently slightly higher than the validation accuracy, as expected. As training progresses, both accuracies stabilize close to 0.90, with the validation accuracy reaching approximately 91.71%. This pattern indicates effective learning without significant signs of overfitting, as shown in FIGURE 10. The close alignment between training and validation accuracy suggests that the model generalizes well to new data, achieving reliable performance. This stable accuracy over many epochs highlights the model's robustness and its capacity to accurately classify data across different sets.

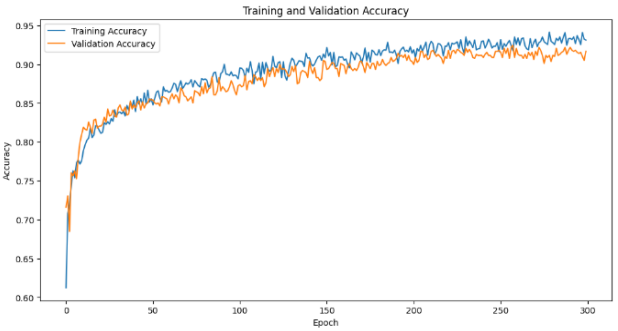

FIGURE 10.Accuracy curves of the VGG16 case in the triple classification.

FIGURE 11 illustrates the reduction in loss across each epoch during the training process of the VGG16 model, showcasing both the training and validation datasets.

Looking at the plot, the training loss is decreasing consistently with epochs meaning that the model can learn from train data well. Initially, validation loss (orange line) also decreases indicating the model generalizes well. But after some point, the validation loss saturates and rises slightly which could be an indication of a little overfitting. VGG16 appears to successfully balance between training loss minimization and generalization, as evidence of this trend.

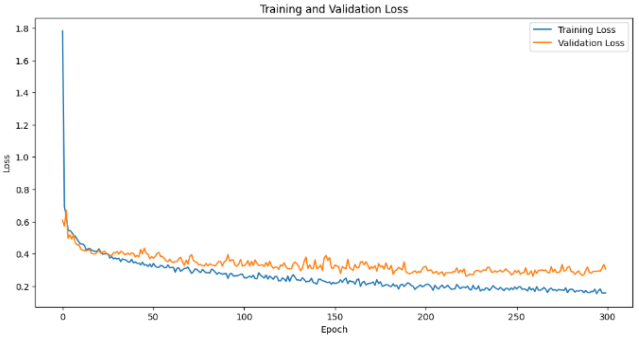

FIGURE 11.Loss curves of the VGG16 case in the triple classification.

The classification report, presented in TABLE 2 ,indicates strong performance of the VGG16 model across all classes in the triple classification task. There is a good balance between precision (0.94) and recall (0.95) for the Low Risk Zone (LRZ) class, resulting in the highest F1-score of 0.95, indicating highly accurate predictions with minimal false negatives. The Moderate Risk Zone (MRZ) class achieved an F1-score of 0.87 with precision at 0.91 and recall at 0.83, showing slightly lower recall in this category. For the High Risk Zone (HRZ) class, an F1-score of 0.92 was achieved, with balanced precision (0.89) and recall (0.96). These metrics confirm that the model is effectively predicting risk levels accurately.

TABLE 2
Precision, Recall, and F1-score for VGG16 in the Triple classification.

	Precision	Recall	F1-Score	Support
MRZ	0.91	0.83	0.87	242

LRZ	0.94	0.95	0.95	301
HRZ	0.89	0.96	0.92	229

The confusion matrix (FIGURE 12) demonstrates that the model accurately classified the MRZ, LRZ and HRZ with accuracies of 83.4%, 95.3%, and 95.0%, respectively. Misclassifications were minimal across categories, with the highest being 10.0% of MRZ instances incorrectly classified as HRZ.

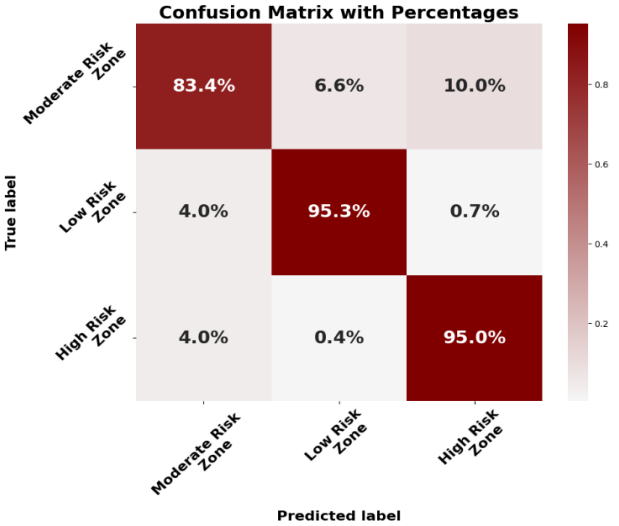

FIGURE 12.Confusion matrix of the validation samples for VGG16 in the triple classification.

2) RESNET50

The accuracy graph over 150 epochs shows a rapid increase in both training and validation accuracy. In a well-trained model, training accuracy should generally remain slightly higher than validation accuracy throughout the training process. Here, the validation accuracy stabilizes at a value close to 95.98%, indicating that the model is generalizing effectively without clear signs of overfitting, as shown in FIGURE 13. The stable accuracy suggests that the model successfully captures the key characteristics of the data, performing well on both the training and validation datasets.

FIGURE 13.Accuracy curves of the ResNet50 in the triple classification.

The loss decrease trend per epoch when training ResNet50 on the training and validation datasets is shown in FIGURE 14. The graph reveals a sharp initial decline in loss, with both training and validation losses converging towards lower values as the number of epochs increases. This steady decrease indicates that the model is effectively learning patterns in the data, resulting in low loss by the end of training. The close alignment between the training and validation loss curves suggests that the model generalizes well, with minimal risk of significant overfitting. Illustrates the reduction in loss (diminution curve) at each epoch during the training of the ResNet50 model on both the training and validation datasets.

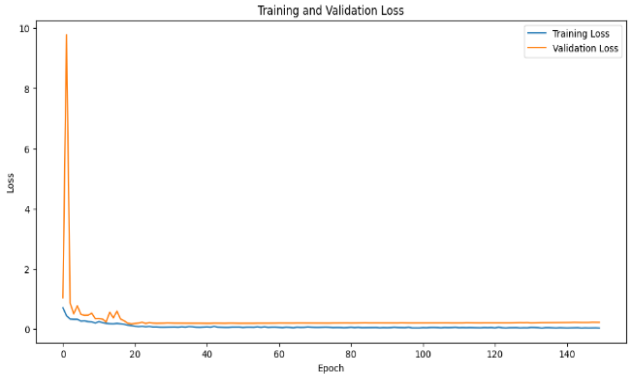

FIGURE 14.Loss curves of theResNet50 case in the triple classification.

The classification report, depicted in TABLE 3 , illustrates an exceptional efficacy across all categories within the model's predictive framework. The Low-Risk Zone (LRZ) category attained the pinnacle F1-score of 0.97, signifying remarkably precise predictions accompanied by a balanced precision (0.97) and recall (0.96). The Moderate Risk Zone (MRZ) category achieved an F1-score of 0.94, with both precision and recall metrics registered at 0.94, demonstrating robust performance within this classification as well. The High-Risk Zone (HRZ) category similarly recorded an F1-score of 0.97, with precision measured at 0.96 and recall at 0.98, reflecting proficient identification of high-risk instances. These performance metrics substantiate the model's credibility in differentiating various levels of risk.

TABLE 3
Precision, Recall, and F1-score for ResNet50 in the triple classification.

	Precision	Recall	F1-Score	Support
MRZ	0.94	0.94	0.94	242
LRZ	0.97	0.96	0.97	301
HRZ	0.96	0.98	0.97	229

The confusion matrix (FIGURE 15) shows accurate categorization for the MRZ, LRZ, and HRZ with accuracies of 94.2%, 96.0%, and 97.8%, respectively. Misclassifications were minimal, with the highest being 3.3% of MRZ instances incorrectly classified as LRZ.

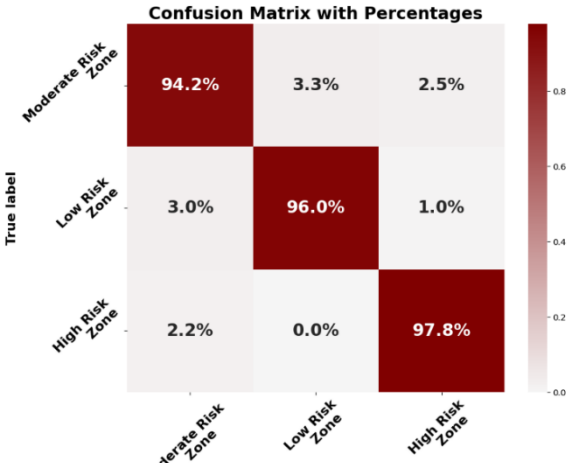

FIGURE 15. Confusion matrix of the validation samples for ResNet50 in the triple classification.

B. QUADRUPLE CLASSIFICATION

1) VGG16

The graph of training and validation accuracy over 150 epochs shows a rapid initial increase, with training accuracy slightly higher than validation, as expected during early learning. Both accuracies stabilize, with validation accuracy reaching around 94.92%, indicating effective learning without overfitting. Minor fluctuations occur later, but both metrics remain closely aligned, reflecting stable performance throughout the training process, as shown in FIGURE 16.

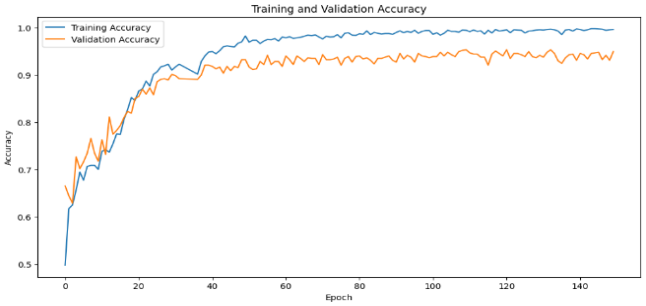

FIGURE 16. Accuracy curves of the VGG16 case in the quadruple classification.

FIGURE 17 delineates the decrement in loss over the course of each epoch during the training regimen of the VGG16 architecture on both the training and validation datasets. The graphical representation exhibits a pronounced reduction in training loss at the outset, which progressively stabilizes as the epochs advance. The validation loss initially follows a similar trend but exhibits some fluctuations, suggesting variations in generalization as training continues. Overall, the loss diminution curve highlights the model's ability to learn effectively from the training data while attempting to maintain accuracy on the validation set.

The classification report, detailed in TABLE 4. , demonstrates robust performance across all four classes. The Low Intermediate Risk Zone (LIRZ) achieved an F1-score of 0.97, indicating high precision and recall. The Low Risk Zone

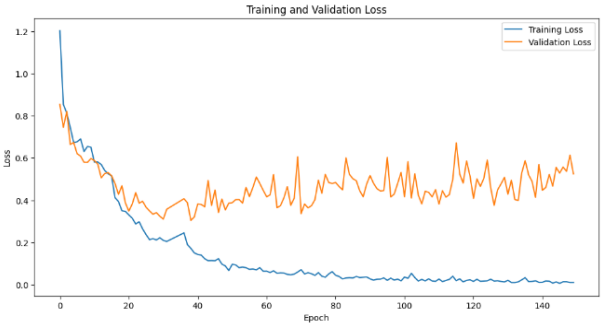

FIGURE 17. Loss curves of the VGG16 case in the quadruple classification.

(LRZ) scored 0.92, reflecting strong predictive accuracy. The High Risk Zone (HRZ) reached an F1-score of 0.96, showcasing effective identification of high-risk cases. The High Intermediate Risk Zone (HIRZ) obtained an F1-score of 0.93, underscoring reliable performance in this category. These metrics confirm the model's capability to accurately classify cases across varying risk levels.

TABLE 4
Precision, Recall, and F1-score for VGG16 in the quadruple classification.

	Precision	Recall	F1-Score	Support
LIRZ	0.89	0.95	0.92	150
LRZ	0.97	0.97	0.97	240
HRZ	0.98	0.95	0.96	228
HIRZ	0.94	0.92	0.93	150

The confusion matrix (FIGURE 18) demonstrates that the model accurately categorized the LIRZ, LRZ, HRZ, and HIRZ with accuracies of 95.4%, 96.7%, 94.7%, and 97.8%, respectively. The highest misclassification rate was 4%, where HIRZ instances were incorrectly classified as LRZ.

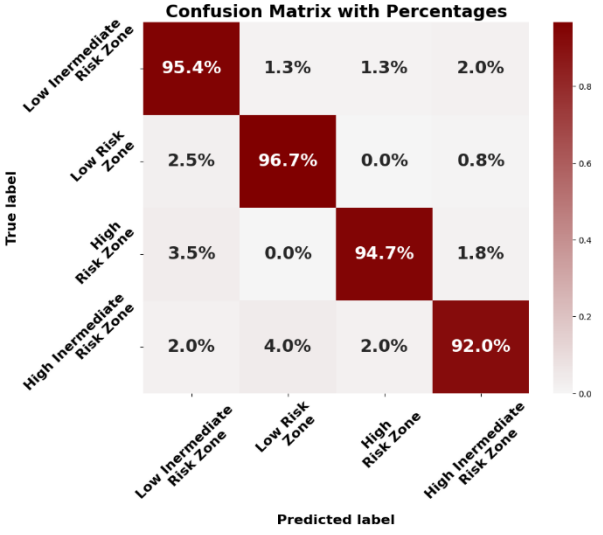

FIGURE 18. Confusion matrix of the validation samples for VGG16 in the quadruple classification.

2) RESNET50

The graph over 150 epochs shows a rapid initial increase in both training and validation accuracy, with training accuracy slightly higher. As training progresses, training accuracy approaches perfect levels, while validation accuracy stabilizes around 94.66%, indicating effective learning without overfitting. Towards the end, both accuracies remain steady, reflecting consistent model performance, as shown in FIGURE 19.

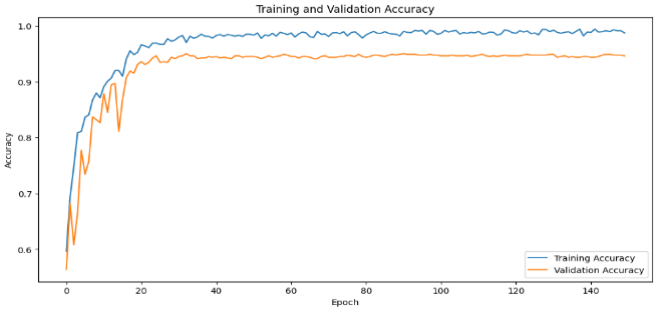

FIGURE 19.Accuracy curves of the ResNet50 in the quadruple classification.

FIGURE 20 illustrates the reduction in loss across each epoch during the training of the ResNet50 architecture on both the training and validation datasets. The graph shows an initial sharp decline in training loss, with the loss values gradually stabilizing as the epochs progress. The validation loss follows a similar trend, although it remains slightly higher than the training loss, indicating good generalization without significant overfitting. This steady decrease in loss demonstrates the model's ability to learn effectively from the data, optimizing its parameters over time for improved performance on both training and validation datasets.

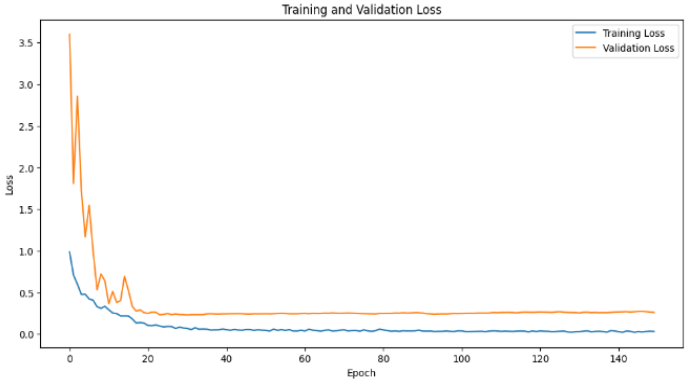

FIGURE 20.Loss curves of theResNet50 case in the quadruple classification.

As is evident from the classification report (Shown in TABLE 5). ResNet50 model performed well across all classes for four fold classification. Among all LRZ class unsurprisingly scored highest F1-score = 0.98, which means that this predicted label is very precise and has a high recall. F1-score of 0.97 came closely behind for the High Risk Zone (HRZ). For the LIRZ, the F1-score is 0.91, and for HIRZ its value attained by the classifier system was 0.89. These metrics confirm the model is effective in matching risk levels with a degree of jaundice severity.

TABLE 5
Precision, Recall, and F1-score for ResNet50 in the quadruple classification.

	Precision	Recall	F1-Score	Support
LIRZ	0.87	0.95	0.91	150
LRZ	0.98	0.98	0.98	240
HRZ	0.98	0.96	0.97	228
HIRZ	0.92	0.87	0.89	150

The confusion matrix (FIGURE 21) demonstrates that the model accurately classified the LIRZ, LRZ, HRZ, and HIRZ with accuracies of 94.6%, 97.9%, 96.5%, and 86.7%, respectively. The most significant error was the misclassification of 8.6% of HIRZ cases as LIRZ.

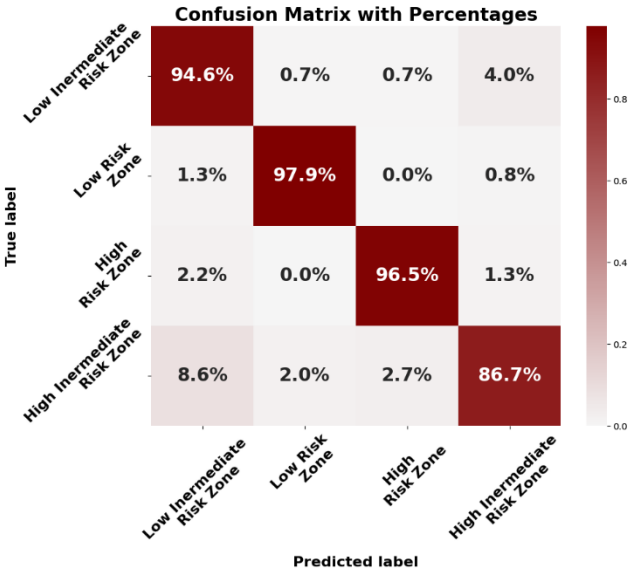

FIGURE 21.Confusion matrix of the validation samples for ResNet50 in the quadruple classification.

V. DISCUSSION

A. TRIPLE CLASSIFICATION

The results indicate that the VGG16 model provides strong classification performance, particularly in the LRZ and HRZ, where the model achieved accuracies exceeding 95%. This suggests that the model is highly reliable in identifying cases requiring immediate attention (high risk) and those that can be safely monitored (low risk). However, there is room for improvement in the MRZ, where 10% of cases were misclassified as "high risk." This indicates that the features distinguishing these two zones are less distinct, and improving the model through additional feature extraction or using larger datasets could enhance performance. The stability of training and validation accuracies, as shown in Figure 9, suggests that the model did not suffer from overfitting, which is a positive outcome. However, minor fluctuations at the end of training

could be addressed by incorporating techniques like early stopping or further hyperparameter tuning.

The ResNet50 model demonstrated exceptional accuracy and reliability, particularly in classifying LRZ and HRZ, with accuracies of 96.0% and 97.8%, respectively. This indicates the model's strong ability to distinguish between risk levels, supporting its use in clinical applications. The model's performance in the MRZ was slightly lower, with 3.3% of cases misclassified as "low risk," possibly due to overlapping features between the two zones. Refining feature extraction or increasing the dataset size could reduce these errors and improve overall performance in this zone. The high accuracy across all zones confirms the model's potential for effective jaundice classification in neonates, making it a valuable tool for clinical settings.

TABLE 6 shows a comparison with the most relevant previous studies that investigated triple classifications of jaundice. Our results outperform those presented in reference [26], which used the VGG16 and ResNet50 algorithms to extract skin features, yielding an accuracy rate of 84.09%. The remaining works listed in the table are all for triple classification, and while some methods outperformed our results, they were more complex than those presented in this work. This highlights the effectiveness and simplicity of the models proposed in this study, providing an efficient and accessible approach for jaundice classification.

Nevertheless, particular constraints may influence the applicability of these results across broader contexts. For instance, the magnitude and breadth of the dataset, especially given its concentration on a singular geographic locale (Mosul), may constrain the efficacy of the model in alternative demographic or geographic settings. Subsequent investigations ought to prioritize the acquisition of more heterogeneous data from varied populations to mitigate this constraint and enhance the model's applicability.

TABLE 6
Comparison with the most related works for triple classification.

Ref.	Data sets	Body Part	Technique	Accuracy
E. P. Rahayu et al	30	Full body	KNN	90
S. Dissaneevate et al	178	Abdomen skin	DT,KNN,and CNN.	96.88
M. S. Jarjees et al	145	Forehead and Abdo-men skin	VGG16, VGG19 ,ResNet50, EfficientNet B0 and B7). DL with fine-tuning.	84.09
Ours	2,563	Skin Region	VGG16, Resnet50	VGG16=91.71 ResNet50=95.98

B. QUADRUPLE CLASSIFICATION

The results from the VGG16 model indicate excellent performance, particularly in the LIRZ and HIRZ, where the F1-scores were 0.97 and 0.93, respectively. The confusion matrix further supports these findings, as the misclassification rates across all zones were minimal, with the highest being only 4% in the HIRZ, misclassified as LRZ. These results suggest that the VGG16 model can effectively differentiate between various risk zones, though there may be room for improvement in further reducing misclassification between similar zones, such as the HIRZ and LRZ. The minor fluctuations in accuracy during the later epochs indicate stability, and the close alignment between training and validation metrics suggests that the model successfully avoided overfitting.Further improvements could be achieved by refining the feature extraction process or increasing the dataset size to provide even greater accuracy and robustness in distinguishing between risk zones.

The ResNet50 model demonstrated robust performance across most classes, particularly in the LRZ and HRZ. where the F1-scores were 0.98 and 0.97, respectively. These high scores indicate the model's reliability in correctly classifying cases that require attention or those that are less concerning. However, the HIRZ presented some challenges, with the F1-score dropping to 0.89 and an 8.6% misclassification rate into the LIRZ. This suggests that the features distinguishing between these two zones are less clear, potentially leading to confusion. Addressing this issue could involve refining the feature extraction process or increasing the size of the dataset for these specific categories. The steady performance observed in both training and validation accuracies, along with the minimal fluctuations towards the end of the training, highlights the model's stability and resistance to overfitting. The overall accuracy across all zones reflects the model's capacity for clinical applications in neonatal jaundice classification, though further improvements in intermediate zones could enhance its precision.

Compared to the results reported by the BiliCam study [33], which achieved a 67% match rate with the Bhutani nomogram, our model demonstrates significantly higher accuracy. Specifically, the VGG16 and ResNet50 models developed in this study attained accuracy rates of 94.92% and 94.66%, respectively, signifying a considerable enhancement in the ability to categorize cases accurately. Additionally, while the BiliCam study reported 19% false negatives and 14% false positives, our model reduced both false negative and false positive rates to 5.08%. This substantial reduction in classification errors underscores the enhanced precision of our approach, making it more reliable for clinical use.

The utilization of these models within clinical environments significantly improves the early identification of neonatal jaundice, thereby facilitating timely interventions and reducing the likelihood of misdiagnosis or unwarranted treatments. The high accuracy achieved in both triple and quadruple classifications provides flexibility in assessing varying degrees of jaundice severity, accommodating different clinical needs. This approach ensures that cases are treated according to their specific risk level, optimizing decision-making in neonatal care and improving patient outcomes. By integrating these models into neonatal care units, healthcare professionals can provide immediate assessments, improving response times and facilitating more precise interventions.

However, it is crucial to safeguard data privacy, especially for sensitive newborn information. Additionally, to avoid algorithmic bias, it is essential to train the models on diverse datasets representing various demographic groups. Continuous monitoring of model performance ensures sustained accuracy and effectiveness in medical diagnostics.

VI. CONCLUSION

The primary objective of this study is to explore non-invasive methodologies for accurately classifying the severity of jaundice in neonates, using skin images as the basis for classification. The research focused on triple and quadruple classifications to assess jaundice severity and guide appropriate treatment strategies. This study contributes to filling a gap in the literature, as research on quadruple classification of jaundice is very limited, making it an important addition to the field. The results demonstrated good classification accuracy, with the VGG16 model achieving 91.71% accuracy in the triple classification and 95.98% using ResNet50. In the quadruple classification, the models achieved 94.92% and 94.66%, respectively.

This study provides a non-invasive and reliable tool for jaundice screening, enhancing early detection and reducing complications, especially in resource-limited environments where healthcare systems face significant challenges. Implementing this method using smartphones or cameras in neonatal care units allows for real-time jaundice assessments and reduces the need for repeated blood tests, with potential adaptation for home care settings.

Future research should aim to expand the dataset by including images from diverse populations, explore alternative deep learning models, and develop a fully automated classification system by integrating cameras into infant incubators to facilitate continuous jaundice monitoring.

ACKNOWLEDGMENT

I would like to thank the management of Ibn Al-Atheer Children's Hospital for their support in making this project successful, as the samples were obtained with their approval. I also thank the Computer Engineering Department for their contribution and supervision of this valuable research.

ETHICAL APPROVAL

This study was conducted in the Mosul city after obtaining parental consent with a full explanation of the study. Additionally, ethical approval was granted by the Research Committee of the Ninawa Health Directorate, Ministry of Health, Mosul, Iraq (Protocol Number: 2023160)

REFERENCES

[1] G. G. Asefa *et al.*, "Determinants of Neonatal Jaundice among Neonates Admitted to Neonatal Intensive Care Unit in Public General Hospitals of Central Zone, Tigray, Northern Ethiopia, 2019: A Case-Control Study," *Biomed Res Int*, vol. 2020, 2020, doi: 10.1155/2020/4743974.

[2] H. Shirzadfar, K. Sheikhi, and Z. Meschian, "The epidemiologic study of neonatal jaundice, relation between jaundice and liver and alternative methods to cure jaundice," *Clinical Practice*, vol. 16, no. 3, 2019, doi: 10.37532/fmcp.2019.16(3).1117-1125.

[3] H. Brits *et al.*, "The prevalence of neonatal jaundice and risk factors in healthy term neonates at National District Hospital in Bloemfontein," *Afr J Prim Health Care Fam Med*, vol. 10, no. 1, 2018, doi: 10.4102/phcfm.v10i1.1582.

[4] M. M. Tiwari and H. N. Pise, "Comparative study between serum and transcutaneous bilirubin measurements in new-borns," *Int J Contemp Pediatrics*, vol. 6, no. 2, p. 817, Feb. 2019, doi: 10.18203/2349-3291.ijcp20190735.

[5] K. Darka and A. Gül, "Correlation of Capillary Tube and Transcutaneous Methods with Serum Biochemistry in Bilirubin Levels in Neonates with Jaundice," *Journal of Contemporary Medicine*, vol. 14, no. 2, pp. 67–71, Mar. 2024, doi: 10.16899/jcm.1434289.

[6] H. S. Khalid and N. A. Fakhre, "SMARTPHONES APP AS PORTABLE COLORIMETRIC SENSOR FOR LOWCOST QUANTIFICATION OF SAFRANIN DYE IN AQUEOUS SOLUTIONS," *Bull Chem Soc Ethiop*, vol. 37, no. 5, pp. 1065–1080, 2023, doi: 10.4314/bcse.v37i5.1.

[7] K. Darka and A. Gül, "Correlation of Capillary Tube and Transcutaneous Methods with Serum Biochemistry in Bilirubin Levels in Neonates with Jaundice," *Journal of Contemporary Medicine*, vol. 14, no. 2, pp. 67–71, Mar. 2024, doi: 10.16899/jcm.1434289.

[8] M. S. Jarjees, S. S. M. Sheet, and B. T. Ahmed, "Leukocytes identification using augmentation and transfer learning based convolution neural network," *Telkomnika (Telecommunication Computing Electronics and Control)*, vol. 20, no. 2, pp. 314–320, 2022, doi: 10.12928/TELKOMNIKA.v20i2.23163.

[9] A. Vuckovic, V. J. F. Gallardo, M. Jarjees, M. Fraser, and M. Purcell, "Prediction of central neuropathic pain in spinal cord injury based on EEG classifier," *Clinical Neurophysiology*, vol. 129, no. 8, pp. 1605–1617, Aug. 2018, doi: 10.1016/j.clinph.2018.04.750.

[10] M. J. Mohammed, E. A. Mohammed, and M. S. Jarjees, "Recognition of multifont English electronic prescribing based on convolution neural network algorithm," *Bio-Algorithms and Med-Systems*, vol. 16, no. 3, 2020, doi: 10.1515/bams-2020-0021.

[11] A. Deshmukh, "Artificial Intelligence in Medical Imaging: Applications of Deep Learning for Disease Detection and Diagnosis," *Universal Research Reports*, vol. 11, no. 3, pp. 31–36, Jun. 2024, doi: 10.36676/urr.v11.i3.1284.

[12] A. Ritahani Ismail, Syed Qamrun Nisa, S. A. Shaharuddin, S. I. Masni, and S. A. Suharudin Amin, "Utilising VGG-16 of Convolutional Neural Network for Medical Image Classification," *International Journal on Perceptive and Cognitive Computing*, vol. 10, no. 1, pp. 113–118, Jan. 2024, doi: 10.31436/ijpcc.v10i1.460.

[13] P. Gayathri, A. Dhavileswarapu, S. Ibrahim, R. Paul, and R. Gupta, "Exploring the Potential of VGG-16 Architecture for Accurate Brain Tumor Detection Using Deep Learning," *Journal of Computers, Mechanical and Management*, vol. 2, no. 2, Jun. 2023, doi: 10.57159/gadl.jcmm.2.2.23056.

[14] A. M. A. BARHOOM, M. R. J. AL-HIEALY, and S. S. ABU-NASER, "BONE ABNORMALITIES DETECTION AND CLASSIFICATION USING DEEP LEARNING-VGG16 ALGORITHM," *J Theor Appl Inf Technol*, vol. 100, no. 20, pp. 6173–6184, Oct. 2022, doi: 10.22214/ijraset.2023.54582.

[15] K. Kamal, H. Ez-Zahraouy, and K. Halloum, "A comparison between the VGG16, VGG19 and ResNet50 architecture frameworks for classiication of normal and CLAHE processed medical images A comparison between the VGG16, VGG19 and ResNet50 architecture frameworks for classification of normal and CLAHE processed medical images," 2023, doi: 10.21203/rs.3.rs-2863523/v1.

[16] N. H. Ali, A. R. Abdullah, N. M. Saad, and A. S. Muda, "Brain cone beam computed tomography image analysis using ResNet50 for collateral circulation classification," *International Journal of Electrical and Computer Engineering*, vol. 13, no. 5, pp. 5843–5852, Oct. 2023, doi: 10.11591/ijece.v13i5.pp5843-5852.

[17] F. Prinzi, T. Currieri, S. Gaglio, and S. Vitabile, "Shallow and deep learning classifiers in medical image analysis," Dec. 01, 2024, *Springer Science and Business Media Deutschland GmbH*. doi: 10.1186/s41747-024-00428-2.

[18] M. N. Mansor *et al.*, "Jaundice in newborn monitoring using color detection method," in *Procedia Engineering*, 2012, pp. 1631–1635. doi: 10.1016/j.proeng.2012.01.185.

[19] B. O. Olusanya, T. M. Slusher, D. O. Imosemi, and A. A. Emokpae, "Maternal detection of neonatal jaundice during birth hospitalization

using a novel two-color icterometer," *PLoS One*, vol. 12, no. 8, Aug. 2017, doi: 10.1371/journal.pone.0183882.

[20] R. Karim, M. Zaman, and W. H. Yong, "A Non-invasive Methods for Neonatal Jaundice Detection and Monitoring to Assess Bilirubin Level: A Review," Jan. 01, 2023, *International Association for Educators and Researchers (IAER)*. doi: 10.33166/AETiC.2023.01.002.

[21] A. Gupta, A. Kumar, and P. Khera, "Method and Model for Jaundice Prediction Through Non-Invasive Bilirubin Detection Technique." [Online]. Available: www.ijert.org

[22] A. Althnian, N. Almanea, and N. Aloboud, "Neonatal jaundice diagnosis using a smartphone camera based on eye, skin, and fused features with transfer learning," *Sensors*, vol. 21, no. 21, Nov. 2021, doi: 10.3390/s21217038.

[23] J. A. Taylor *et al.*, "Use of a Smartphone App to Assess Neonatal Jaundice."

[24] S. Swarna, S. Pasupathy, B. Chinnasami, N. M. D., and B. Ramraj, "The smart phone study: assessing the reliability and accuracy of neonatal jaundice measurement using smart phone application," *Int J Contemp Pediatrics*, vol. 5, no. 2, p. 285, Feb. 2018, doi: 10.18203/2349-3291.ijcp20175928.

[25] J. Castro-Ramos, C. Toxqui-Quitl, F. Villa Manriquez, E. Orozco-Guillen, A. Padilla-Vivanco, and JJ. Sánchez-Escobar, "Detecting jaundice by using digital image processing," in *Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing XXI*, SPIE, Mar. 2014, p. 89491U. doi: 10.1117/12.2041354.

[26] W. Y. Hsu and H. C. Cheng, "A fast and effective system for detection of neonatal jaundice with a dynamic threshold white balance algorithm," *Healthcare (Switzerland)*, vol. 9, no. 8, Aug. 2021, doi: 10.3390/healthcare9081052.

[27] M. Aydın, F. Hardalaç, B. Ural, and S. Karap, "Neonatal Jaundice Detection System," *J Med Syst*, vol. 40, no. 7, Jul. 2016, doi: 10.1007/s10916-016-0523-4.

[28] R. Angelico *et al.*, "A novel mobile phone application for infant stool color recognition: An easy and effective tool to identify acholic stools in newborns," *J Med Screen*, vol. 28, no. 3, pp. 230–237, Sep. 2021, doi: 10.1177/0969141320974413.

[29] A. Yaseen Abdulrazzak, S. Latif Mohammed, A. Al-Naji, and J. Chahl, "Computer-Aid System for Automated JaDetectionundice," *Journal of Techniques*, vol. 5, no. 1, pp. 8–15, Mar. 2023, doi: 10.51173/jt.v5i1.1128.

[30] E. P. Rahayu, M. N. Widyawati, and S. Suryono, "Euclidean distance digital image processing for jaundice detect," *IOP Conf Ser Mater Sci Eng*, vol. 1108, no. 1, p. 012022, Mar. 2021, doi: 10.1088/1757-899x/1108/1/012022.

[31] S. Dissaneevate *et al.*, "A Mobile Computer-Aided Diagnosis of Neonatal Hyperbilirubinemia using Digital Image Processing and Machine Learning Techniques," *International Journal of Innovative Research and Scientific Studies*, vol. 5, no. 1, pp. 10–17, 2022, doi: 10.53894/ijirss.v5i1.334.

[32] M. S. Jarjees, N. Abdulkadir, and K. Kandla, "Neonatal Jaundice Detection System Using Convolutional Neural Network Algorithm," 2023, doi: 10.14704/nq.2022.20.8.NQ44777.

[33] L. De Greef *et al.*, "BiliCam: Using mobile phones to monitor newborn jaundice," in *UbiComp 2014 - Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, Association for Computing Machinery, Inc, 2014, pp. 331–342. doi: 10.1145/2632048.2632076.

[34] B. J. Lauer and N. D. Spector, "Hyperbilirubinemia in the newborn," Aug. 2011. doi: 10.1542/pir.32-8-341.

AUTHORS BIOGRAPHY

Banan Khalid Abdulkader is a master's student in computer engineering at the University of Mosul, Iraq. She earned her bachelor's degree in engineering from the same university in 2021. Throughout her academic journey, she has gained significant expertise in machine learning and artificial intelligence, with a special focus on developing machine learning models for data analysis and accurate prediction of outcomes. Her skills include building models using deep learning techniques and neural networks, enabling her to offer effective solutions. In addition to her studies, she continues to research how to leverage these technologies to develop innovative solutions in areas like data analysis and AI, which enhances her academic and professional prospects.

Mazin H. Aziz earned his Bachelor and Master of Science degrees in Electrical Engineering, specializing in Electronics and Communication, from the University of Mosul in 1989 and 1993. In 2019, he completed his PhD in Electrical Engineering, focusing on electronics, also at the University of Mosul. Since 2008, he has been a lecturer in the Computer Engineering Department at the College of Engineering, University of Mosul. Before this, he worked at the Al Kindy Research and Development Center, specializing in digital electronics and computer systems. His research interests encompass image processing, computer vision, artificial intelligence, and biomedical engineering. He has contributed to many publications and serves as a reviewer for several reputed journals and conferences.mazin.haziz@uomosul.edu.iq