Original Research Article

Low cost ML based collaborative approach for baby monitoring

Abstract. There are numerous baby-monitoring devices available in the market that parents use to keep an eye on babies while they are away. The majority of them are reliant on the installation of expensive hardware, which many parents cannot afford. Another issue with these devices is that they detect high-pitched sounds and frequently give false alarms, causing both children and parents to be disturbed. The majority of smartphone applications in the market work on sound wave and only sound an alarm when the infant start crying. In this project, we proposed the design of a mobile application to detect the status of a baby inside a crib/ on a bed. The application will alert parents when their child requires assistance, will be able to determine whether the child is sleeping in a safe or hazardous position, and will keep track of the child's sleeping patterns. It is less reliant on hardware, making it less expensive. Here the only requirement is two paired mobile phones with the application installed instead of expensive hardware (IoT-based devices). The application is utilizing the transfer-learning technique on tensor flow lite Mobilenet classification and SSD_mobilenet_V1_coco object detection models. The accuracy of the model is 97% for the Mobilenet classification model and 98% for the object detection model.

Keywords: Baby-monitoring devices ,Transfer-learning, Mobilenet classification, Object detection.

1 Introduction

A new born brings lots of happiness in the family, but with plethora of responsibilities. There are many questions, which can be stressful for new parents. Like is it necessary for someone to be constantly there with the baby, what if the baby needs attention and there is no one available to provide it. In addition to babysitting, a regular daily routine/office works make it impossible to be there with the babies at all the times. Continuously monitoring a baby is a difficult task, especially for working parents, carrying out their babies all the time is not possible [1].

The majority of the parents have installed CCTV cameras; however, these are unable to inform them in the event of an emergency. Because of the electromagnetic waves emitted by various wearable devices, many parents do not like these devices and the other problem with these devices is that they may disturb a baby's sleep. All other IoT-based devices are costly because of the expensive hardware dependency. These devices are not affordable to many of the parents. Almost all of these devices detect high pitch sounds & many a time give false alarms. Some mobile applications are available, but the problem with all of these devices is that they are based on the sound that alerts parents only when the baby is crying. **Fig. 1.** Shows different baby monitoring system available in the market.

Fig. 1. Baby-monitoring system available in the market

According to Stanford Children's Health Organization, babies use to sleep most of the time but they sleep in short segments and they need undisturbed sleep. American Academy of Pediatrics (AAP) suggest babies should never sleep on their side or on stomach until they are one year old. Infants should not sleep on the side either as they can easily roll to their tummies. The tummy sleeping and side sleeping are considered unsafe or hazardous for babies as they are not able to take sufficient oxygen in these positions. Pediatricians recommend laying them only on their back through the first year of the babies. It increases the airflow in the infants. This can reduce the risk for SIDS(unexpected death generally occurred during sleep, and happens with healthy young infant under the age of one)[2].**Fig. 2.**Shows the babies sleeping in Hazardous position vs Safe position.

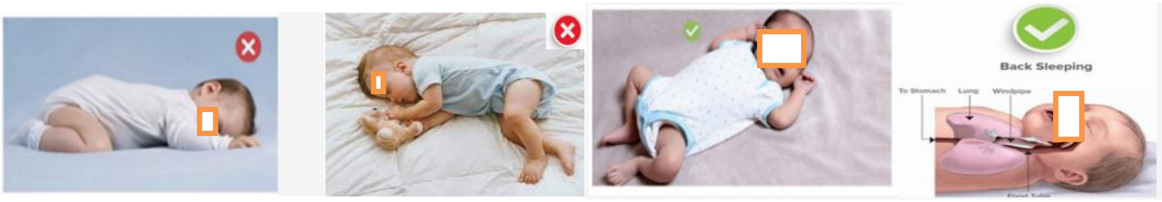

Fig. 2. Hazardous position vs Safe position

When the babies are sleeping, they commonly slip out of the bed. Sometimes it may cause head, neck, or body pain in babies[3]. Keeping all these points in consideration, monitoring babies during their sleep is required and should be affordable to all parents.

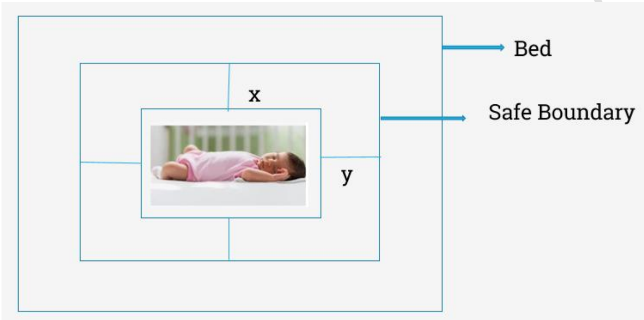

Fig. 3. Safe boundary for babies

In this paper we have defined the safe boundary for babies which is 25% inside from the bed edges.If a baby cross this boundary,it means they are near to bed edge and outside the safe zone which needs parents attention.we have calculated distance x and y from the boundary edges using open cv and boundary boxes coordinates defined by object detection model.

This research proposes an algorithm to monitor children during their sleep based on the status and position of babies on the bed / inside the crib and when detecting any wrong pattern the smart monitoring system alerts the parents by calling on their mobile phone. The design of a mobile application has been suggested in this studyusing transfer-learning techniques on state-of-the-art CNNs to identify a child's status from visual monitoring. Pre-trained models have been selected based on low memory requirements with high accuracy, recognizing the finite compute/memory availability in mobile-based applications. TensorFlow lite Mobilenet classification model is used to check the status of the baby on the bed/inside the crib while SSD MobileNet V1 coco model is used to detect the position of a baby inside the crib/on the bed.

1.1 Challenges

While CNN's are known to be able to achieve high accuracy rates providing enough data for supervised image classification, in practice, the resources needed to collect large amounts of videos and images of babies is a hurdle for commercialization. It is difficult to find a large number of parents who will offer video streams of their child in the crib/on bed given the privacy concerns[4]. A total of three hundred images have been collected for this project from the internet.

Model architecture selection and optimization.Multiple model designs such as Efficient Net, Inception-V3, ResNet50, and Mobilenet have been tested to determine the greatest feasible accuracy/precision. To understand the impact of transfer learning, Pre-trained model weights have been fine-tuned by freezing and unfreezing different layers of weights during training. Furthermore, as an edge computing application where the inference step will be processed on a mobile device with memory, computation, and power consumption limitations, TensorFlow lite Mobilenet for classification and SSD Mobilenet model for object detection have been chosen with a minor loss of accuracy in exchange for increased speed with limited resource requirement.

Understanding and overcoming domain challenges. To be truly viable, an application should have the ability to accurately classify images captured from unseen targets. For example, there can be significant domain variation caused by different parameters such as different settings of mobile phone cameras; various bed/crib designs, different baby sizes, different lighting circumstances, and several types of toys in the crib or on the bed. This study tried to understand the domain adaptability of the trained model by testing the model against datasets with some domain changes using augmentation techniques (like random image cropping, flipping, and rotation).

2 Related works

It is simple to install and apply a mobile application in the real world. It eliminates the requirements of a wearable device or camera designed specifically for software, and because of this, it is economical and does not rouse sleeping infants [5].

A paper published by Choi, Soohyun, and his team illustrate the design of an automated baby monitoring system "CCTV Beebe(CCBeBe)" that uses AI to check the infant's posture during sleep and crying. Their objective was to monitor dangerous lying posture, infant cry, and video streaming of baby, sound transmission, acknowledging the baby's emotion, and alerting the parents about the changing events. If a cry is acknowledged, the application sends an alert message to parents' mobile. They have implemented "Open Pose, EfficientNet, WebRTC, and Facial-Expression-Recognition Pytorch"[5][6]. This posture detector model had a lower accuracy. In this study, they did not consider the baby's distance from the bed edge.

In June 2017, Lily Cheng from Stanford University published her research on identifying a child's status through a visual monitoring gadget. Three different models that include ResNet18, AlexNet, and SqueezeNet were trained using five pre-defined classes. Despite being a significantly smaller model, SqueezeNet was able to deliver a test accuracy of 94.6%.This research shows that in a real-life application, smaller model is preferable due to efficient model size. The model has shown to have low domain adaptation when evaluated with video frames acquired using a different camera position or discovered online[4]. More work is required to increase the model's ability to categorize images acquired from unknown domains

In 2018, M. P. Joshi and D. C. Mehetre proposed a Smart Cradle System based on IoT for Baby Monitoring with an Android Application. When the application detect a baby cry, the smart cradle swings automatically. In addition, if the baby continues to cry for a specific amount of time, a buzzer sounds and an alert and call to parents' mobile. The system designed by them is cost-effective and simple to use, making it ideal for busy parents. In the future work, they suggested more features, like an IR (Infrared) camera for night vision and other client programs, such as those for iOS can be designed[7].

Yogita K. Dubey and Sachin Damkein proposed an application using image processing to ensure the baby's safety and activity monitored by their busy parents. The technology recognizes the baby's movement, crying, and current position. If an aberrant action is identified, the system sends an email to the user with a notification in the form of text and images of the baby. They used Haar cascade face Detection method to construct the software for the required outcomes, with the Python language used to design the complete code by utilizing DIP techniques through the OpenCV library. It is able to determine the position of the baby on the bed using the Raspberry Pi Camera [8].

In 2020, Nour Mahmoud & his team from Islamic University Al-Madinah Saudi Arabia worked on "A Cost-effective IoT based system for monitoring baby incidents by deaf parents". Their findings suggest an algorithm for keeping track of a baby's health as and when he or she sleeps. The smart monitoring device alerts mothers by shaking a wrist bracelet or ringing their mobile if there is some surroundings change. The proposed system has been built and tested, and it is superior to competing systems in terms of cost of implementation, memory consumption, and alert accuracy[9]. This IoT-based device is costly.

It is not easy to create a small but efficient neural network. The challenging task is to minimize the number of model parameters while maintaining comparable accuracy and making training faster. Lightweight networks such as SqueezeNet, MobileNet, ShuffleNet, and ESPNet produce promising results not only for image classification but also for object detection. [10]

According to Mark Sandler's "MobileNetV2: Inverted Residuals and Linear Bottlenecks" paper, a new neural network architecture can be optimized for mobile environments and limited resource availability. Through a significant reduction in operations and memory, in their network, mobile computer vision models are made to be tailored to mobile devices, while retaining the same accuracy[11].

3 Background and Research contributions

It is crucial to keep an eye on babies, but it is difficult for parents to do so all of the time. Even though most of the parents use CCTV cameras to track their babies, this is unsuccessful in alerting them in the event of an emergency. Some wearable gadgets are available that push alert signals but most of the parents do not want to use these because of the radiation emitted by these systems. These devices also necessitate the purchase of costly specialized equipment[5][6].

The goal of this research is to create a mobile application that can identify a baby's position in a crib or bed. The application should notify the parents when the baby needs them. It should be able to detect whether the baby is sleeping in a safe position or not. The proposed mobile application would track the sleeping habits of the baby. There should be no hardware dependency, which make it affordable to everyone.

4 Research Methodology

Crisp-DM approach has been selected to implement this project. CRISP-DM (Cross Industry Standard Process for Data Mining) is a methodology developed in 1996 to help shape Data Mining projects. It comprises of six processes for conceiving a Data Mining project, with cycle iterations based on the needs of developers. The six phases are Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, and Deployment.

The market for smart baby monitoring devices is predicted to increase at a compound annual growth rate(CAGR) of 8.0 percent from FY2020 to FY2028, from US$ 989.38 million in 2020 to US$ 1,815.31 million in 2028[12][13]. **Fig. 4.** shows the market of smart baby monitoring devices.

Fig. 4. Smart baby monitor market growth

The market for smart baby monitors is growing due to rise in the number of working mothers around the world. The popularity of smart baby monitoring gadgets is increasing due to the growing number of nuclear families and working parents around the world. Working parents are actively seeking nannies or childcare centers to look after their children while they are away from home. With the increasing concerns associated with child safety, parents are more likely to invest in smart baby monitoring systems.

In developed countries as well as in developing countries, like United States, China, India, and Canada, the number of working mothers is on the rise. Smart baby monitoring solutions assist parents in carrying out their duties effectively and stress-free. These solutions enable parents and their children to communicate in real-time via cell phones. As a result, market growth is aided by the growing use of smartphones, which parents use to install apps to keep track of their children[13].

The market for baby monitoring system is increasing, due to global electronics industry growth and the advent of digitalization. Additionally, the rise in both parents working households and especially the number of working mothers are two important reasons driving the baby monitoring system demand. However, the expensive cameras, particularly in developing nation, affects the industry expansion. This also raise product storing costs to fulfill the need for AI-enabled monitors and more required features, which raises the price of the product as well as complete cost of production. Over the last few years, the demand for video baby monitors also increased among working parents in LAMEA and Asia-Pacific countries due to significant growth of economy[13].

Despite significant progress in lowering infant mortality rates around the world, WHO (World Health Organization) estimates show that approximately 2.4 million newborns died in the first month following birth in 2019. Furthermore, the WHO reported that approximately 7,000 newborns died every day in 2019, accounting for 47 percent of all children under the age of five. In 2019, Asian countries such as India had a neonatal death rate of 21.7 per 1,000 live births, much higher than the global average of 17.5 per 1,000 newborns reported in 2019. As a result, the demand for enhanced newborn care and monitoring goods is likely to rise in the next years, thereby boosting the global market for baby monitoring systems. The number of new generation parents who are aware of IoT baby monitoring systems has risen dramatically. The expensive cost of installing infant monitoring devices may limit the market's expansion[14].

In this research, 300 image samples of different babies on the bed and inside the crib were selected from the internet. All of the images are in jpg format, with file size ranging from 3-5 KB. A large number of images could not be collected due to resource constraints. In practice, a large number of images must be collected and properly labelled to achieve the highest accurately classify images drawn from unknown target domains. Varied parameters, such as different settings of a mobile phone camera, different bed/crib designs, and varying sizes of babies, different lighting conditions, and different types of toys inside the crib, can induce significant domain shift. This reserach attempted to evaluate the trained model's domain adaptability by evaluating it against datasets that had some domain change with augmentation techniques (such as random image cropping, flipping, and rotation). By using data augmentation a total of 3500 images has been generated and used to fine-tune custom train the classification and object detection pre-trained model.

For this project, we have collected 300 images from the internet. All the images has been converted into squares of 224x224 pixels. Then all the images were manually classified into different categories. Dataset was further split at random into training, validation and test set at a ratio of 80:10:10.As data size was too small to accurately classifying images in different categories, we have generated more images with augmentation technique using different parameters before feeding into different models. Only training images has been augmented here. For object detection model, labelling of image has been done using labelimg tool.

Fig. 5. Augmented images:

Fig. 5. shows the sample images generated by augmentation technique. For each image, 20 new images have been generated by zooming in, zooming out, shifting, rotating and applying many other editing operations. The training data size has been increased from 175 to 3500.

Getting plethora of images was quiet challenging to train any convolutional neural network from scratch. Moreover, there are various state of art pre-trained models available for image classification and object detection. Which has been used with Tensor flow lite model makers to make it usable for mobile application. The TensorFlow Lite Model Maker library simplifies the process of training a TensorFlow Lite model using custom dataset. It uses transfer learning to reduce the amount of training data required and shorten the training time. The Model Maker library currently supports Image Classification, Object detection, Text Classification, BERT Question Answer, Audio Classification and Recommendation based on the context information for on-device scenario.

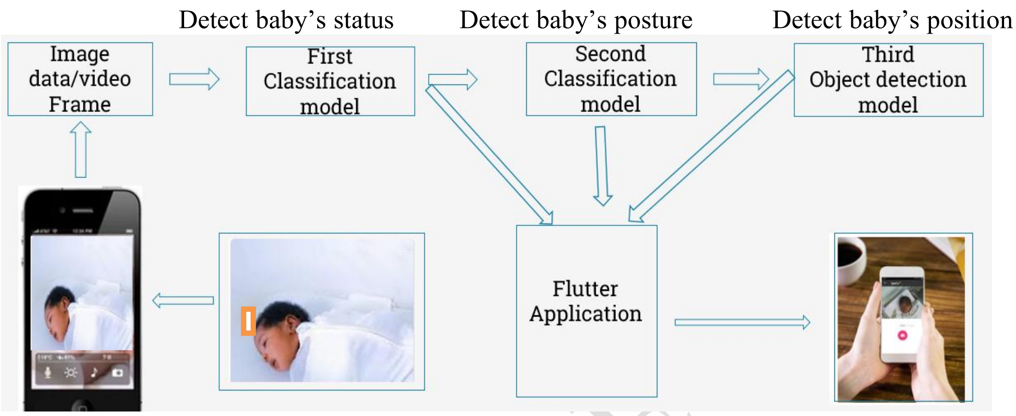

Fig. 6. Process flow diagram for baby monitoring system

Fig. 6.shows two MobileNet images classification models and one SSD(Single Shot Detection) object detection model were used in this project .The first image classification model will determine whether the baby is sleeping or not. If baby is not sleeping, the model will alert the parents by calling on their mobile phones. If the baby is sleeping, then the second model will be triggered. This model will examine whether the baby is in a safe sleeping position or not.

This model will generate an alert and call parents if the baby is not sleeping in a safe position. If the baby is sleeping in a safe position, the third model will be triggered. This object detection model will check whether the baby is sleeping in a safe area or not. If the baby is sleeping outside the safe boundary, it will give a call on parents' mobile phone. When the baby is sleeping within a safe boundary, it will keep checking until there is a change in posture or position. Above Process flow diagram of overall proposed application, illustrate the functionality of the application.

We evaluated the model with actual images as well as augmented images. With actual images, the test accuracy was as low as 83% only but with augmented images fine-tune Mobilenet_V2_classification test accuracy was 97% for first classification model that was used to detect whether baby is sleeping or not.

Images	Count	Accuracy Train%	Accuracy validation%	Accuracy Test%
Actual	175	100	96.8	83.3
Augmented	3500	100	98.8	**97.0**

Table 1. Accuracy obtained with Mobilenet V2 classification model (baby status)

Table 1 shows the accuracy comparison of first model to detect whether the baby is sleeping or not. With actual images (175 images), test accuracy is 83.3%.While with augmented images (3500 in count), test accuracy has increased to 97%.

Images	Count	Accuracy Train%	Accuracy validation%	Accuracy Test%
Actual	88	100	88.9	77.8
Augmented	2262	100	97.3	**96.0**

Table 2. Accuracy obtained with Mobilenet V2 classification model (baby Pose)

Table 2 shows the accuracy comparison of first model to detect whether the baby is sleeping in safe position or not. With actual images (88 images), test accuracy is 77.8%.While with augmented images (2262 in count), test accuracy has increased to 96%.

Object_detection_SSD_Mobilenet_V1 has shown 95-98% confidence interval for predicting babies as a person. An alert signal has been sent to the parent’s mobile, whenever the baby is either not sleeping or sleeping in a dangerous position or outside the safe zone.

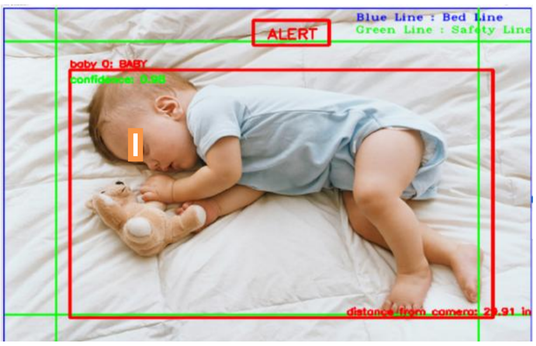

Fig. 7. Baby *sleeping in dangerous pose*

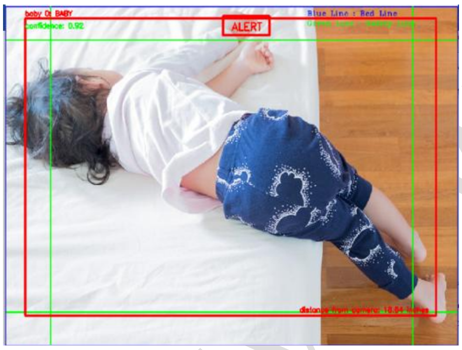

Fig. 8. Baby sleeping outside the safe region

Fig. 7.shows model prediction of baby sleeping in dangerous position and **Fig. 8.** shows model correctly detected baby sleeping outside the safe boundary. In both the cases, the model has generated alert.

Mobile application has been built using flutter app. Flutter is Google's UI toolkit for building beautiful, natively compiled applications for mobile, web, desktop, and embedded devices from a single codebase. This app named as baby monitoring system can be installed on any android phone and after filling minimum details required it can be used to monitor a baby on bed or crib.

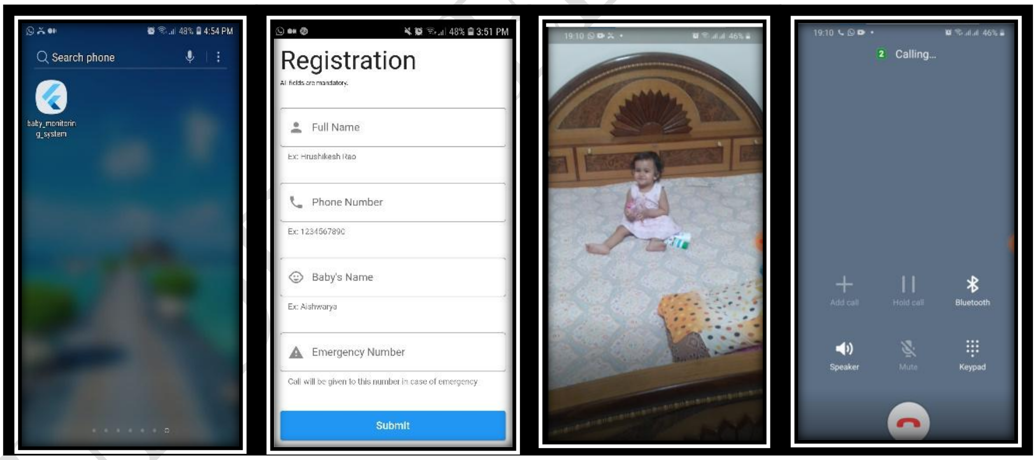

Fig. 9.shows how the application is working on mobile. As the baby is not sleeping, a call has been sent to parents mobile.

Conclusions and Future Scope

A smart and cost effective solution was built using transfer learning techniques on state of art CNN pre trained model for classification and object detection. Model is working quite well with different unseen scenario. All the images taken is in good lighting condition, it cannot perform well in dark room or poor lighting condition. In future model can be designed in keeping this point in consideration.

The MobileNet class of convolutional neural networks work better and are far smaller than other popular models used for classification as well as for object detection. With Mobilenet V1, we got an accuracy as high as 100% with training data (175 sample images) without augmentation, and 93.5% with validation dataset. However, the model performed poorly with the classification task with the new set of test data. The reason behind this was less number of samples and unbalanced data.

By applying the Augmentation technique, the sample size has been increased. With this big dataset, validation accuracy has increased to 97.5% while on test data accuracy is 97%, which is quite impressive. However, we checked the performance of the model against other lightweight models as well. With EfficientNet-Lite0, validation accuracy is 97.8%, and test accuracy was 95.8% only. With MobileNetV2, the validation accuracy has increased to 98.8%; test accuracy has increased to 97.3%. With Inception V3, validation accuracy is 97.5%, and test accuracy is 97.0% only.

Model Name	Train Accuracy	Validation Accuracy	Test Accuracy
EfficientNet-Lite0	97.8%	97.5%	95.8%
Mobilenet V2	98.8%	98.8%	97.3%
Inception V3	96.5%	97.5%	97.0%

Table 3. Accuracy comparison with three different models

Table 3 shows the accuracy achieved by three different models. Mobilenet V2 has somewhat higher accuracy than other lightweight models. This model is faster in the mobile application as per TensorFlow model Zoo GitHub description. Hence based on these data, we have selected MobileNet V2 model for the classification task.

To detect baby's position within the safe boundary zone, we have used SSD_Mobilenet_V1 model, which can detect a baby as a person with a high confidence score of 95-98%.We have not used bigger models YOLO or Faster RCNN as these requires more resources and not suitable for mobile application.

New scenarios can be taken care of in the future work. All of the images taken are in good lighting; they might not perform well in a dark room or poor lighting condition. Future model can be designed keeping this point in consideration. More features can be added for night vision. In future, color and brightness adjustments can be considered to synthesize lighting scenarios.

References

[1] A. F. Symon, N. Hassan, H. Rashid, I. U. Ahmed, and S. M. T. Reza, "Design and development of a smart baby monitoring system based on Raspberry Pi and Pi camera," *4th Int. Conf. Adv. Electr. Eng. ICAEE 2017*, vol. 2018-Janua, pp. 117–122, 2017, doi: 10.1109/ICAEE.2017.8255338.

[2] "www.stanfordchildrens.org," 2019. https://www.stanfordchildrens.org/en/topic/default?id=infant-sleep-90-P02237.

[3] "www.medicalnewstoday.com," 2020. https://www.medicalnewstoday.com/articles/322208.

[4] L. C. June, "Robo-Nanny : ConvNets for Intelligent Baby Monitoring," 2017.

[5] S. Choi, S. Yun, and B. Ahn, "Implementation of automated baby monitoring: CCBeBe," *Sustain.*, vol. 12, no. 6, 2020, doi: 10.3390/su12062513.

[6] "www.mdpi.com," 2020, [Online]. Available: https://www.mdpi.com/2071-1050/12/6/2513.

[7] M. P. Joshi and D. C. Mehetre, "IoT Based Smart Cradle System with an Android App for Baby Monitoring," *2017 Int. Conf. Comput. Commun. Control Autom. ICCUBEA 2017*, pp. 1–4, 2018, doi: 10.1109/ICCUBEA.2017.8463676.

[8] S. gajanan D. Yogita Dubey, "Baby monitoring system using image processing and IoT," *Int. J. Eng. Adv. Technol.*, vol. 8, no. 6, pp. 4961–4964, 2019, doi: 10.35940/ijeat.F9254.088619.

[9] N. Mahmoud, A. Namoun, Ahmad B, Sami S. Albouq, and A. Ahmed, "A Cost effective IoT based system for monitoring baby incidents by deaf parents," 2020.

[10] H. Phan, D. Huynh, Y. He, M. Savvides, and Z. Shen, "MoBiNet: A mobile binary network for image classification," *Proc. - 2020 IEEE Winter Conf. Appl. Comput. Vision, WACV 2020*, pp. 3442–3451, 2020, doi: 10.1109/WACV45572.2020.9093444.

[11] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 4510–4520, 2018, doi: 10.1109/CVPR.2018.00474.

[12] "smart-baby-monitor-market-forecast-to-2028.pdf." [Online]. Available: https://www.researchandmarkets.com/reports/5313989/smart-baby-monitor-market-forecast-to-2028.

[13] "www.alliedmarketresearch.com," 2020. https://www.alliedmarketresearch.com/baby-monitor-market.

[14] Mr.Abhishek Paliwal, "www.researchdive.com," 2020. https://www.researchdive.com/press-release/baby-monitor-market.html. (accessed Jul. 24, 2021).