Photoplethysmography Signal Quality Assessment using Attentive-CNN Models

Leonardo Silva[1], **Rafael Lima**[1], **Giovani Lucafo**[1], **Italo Sandoval**[1], **Pedro Garcia Freitas**[2] and **Otávio A. B. Penatti**[1]

[1]Samsung R&D Institute Brazil (SRBR), Campinas, Brazil

[2]Department of Computer Science – University of Brasília, Brasília, Brazil.

Abstract. *Due to the rapid popularization of wearable computers such as smartwatches, Health Monitoring Applications (HMA) are becoming increasingly popular because of their capability to track different health indicators, including sleep patterns, heart rate, and activity tracking movements. These applications usually employ Photoplethysmography (PPG) sensors to monitor various aspects of an individual's health and well-being. PPG is a non-invasive and cost-effective optical technique based on the detection of blood volume changes in the microvascular bed of tissue, capturing the dynamic physiological changes in the body with continuous measurements taken over time. Analyzing PPG as a time series enables the extraction of meaningful information about cardiovascular health and other physiological parameters, such as Heart Rate Variability (HRV), Peripheral Oxygen Saturation (SpO2), and sleep status. To enable reliable health indicators, it is important to have robustly sampled PPG signals. However, in practice, the PPG signal is often corrupted with different types of noise and artifacts due to motion, especially in scenarios where wearables are used. Therefore, Signal Quality Assessment (SQA) plays a fundamental role in determining the reliability of a given PPG for use in HMA. Considering this, in this work, we propose a novel PPG SQA method focused on the balance between storage size and classifier quality, aiming to achieve a lightweight and robust model. This model is developed using recent advances in attention-based strategies to significantly improve the performance of purely Convolutional Neural Network (CNN)-based SQA classifiers.*

1. Introduction

The increasing impact of continuous health monitoring on the market of smart wearable devices, particularly smartwatches, is noteworthy. The health monitoring applications integrated into these devices are increasingly used for estimating various physiological factors from users, including Heart Rate Variability (HRV), Peripheral Oxygen Saturation (SpO2), and sleep quality. Photoplethysmography (PPG) signal serves as a central component for estimation in these applications, due to its non-invasiveness and cost-effective on-device operation. Smart wearable devices, equipped with PPG sensors, have gained a widespread adoption, ranging from fitness bands to smart rings. According to [Cisco 2019], the global surge in wearable usage is evident, with estimates projecting over 1 billion wearables worldwide by 2022. Additionally, [Gartner 2021] shows an annual spending surpassing U$80 billion globally. These devices offer the flexibility to customize and incorporate various sensor types, transmission units, and remote computing resources, providing end-to-end health solutions to users. Within this context, PPG emerges as a convenient technique, playing a substantial role in wearable health monitoring systems. It continuously delivers physiological parameters that can be used to estimate health information, including heart rate, respiratory rate, and oxygen saturation, contributing to the comprehensive nature of these health-focused wearables.

PPG operates by emitting a luminous signal onto the surface of the user's skin and capturing its reflection/transmission, which variate proportionally to the blood volume in the tissue. This measurement provides valuable information about the cardiovascular system. One challenge affecting the on-device performance of PPG is its vulnerability to noise, including motion artifacts [Chatterjee et al. 2022], which can distort the signal's morphological properties and lead to erroneous estimation of the physiological variables. Due to the potential life-threatening consequences associated with erroneous assessments derived from these signals, such unreliable performance is unacceptable in real-world applications. The presence of noisy signal fragments is the primary motivation behind the development of Signal Quality Assessment (SQA) techniques. This is essential to prevent misinterpretation by discerning between dependable and undependable segments in PPG. In other words, to enhance the reliability of such application, a Signal Quality Classifier (SQC) step is typically employed, allowing the distinction of high-quality signal segments, suitable for variable estimation, from low-quality ones.

Due to the importance of the task of classifying signal segments into reliable or unreliable, some researchers have invested valuable efforts in developing classification techniques for signal quality assessment, as reviewed by [Gambarotta et al. 2016]. For instance, [Elgendi 2016] suggested a technique utilizing indices to assess quality. [Selvaraj et al. 2011] introduced a statistical method that involves calculating kurtosis and Shannon Entropy to identify motion artifacts and noise in PPG data. [Li and Clifford 2012] introduced an alternative statistical approach, utilizing dynamic time warping to elongate each heartbeat for alignment with a dynamic template. This method integrates various features associated with signal quality. In the classification phase of Li & Clifford's approach, a multi-layer perceptron is employed to understand the correlations among parameters in the context of both high- and low-quality pulses. [Sun et al. 2012] introduced an approach that utilizes the morphological features of the signal for evaluating its quality. [Li et al. 2011] identified four waveform characteristics to evaluate signal quality through the application of a decision tree. Likewise, [Sukor et al. 2011] utilized a basic decision-tree classifier to determine, with pulse-by-pulse precision, whether a specific pulse is suitable for use or not. [Naeini et al. 2019] introduced one of the initial methods based on Machine Learning (ML) to categorize the signal into 'reliable' or 'unreliable' categories. In a recent development, [Freitas et al. 2023b, Freitas et al. 2023a] introduced a SQA technique that transforms PPG signals into bi-dimensional representations and subsequently employs a vision transformer to evaluate their quality.

[Lucafo et al. 2022] proposed a categorization of SQA techniques in "rule-based" and "learning-based" methods. This categorization is general enough to distinguish between the most of the aforementioned papers in the literature. Other taxonomies, such as the one proposed by [Zhang et al. 2022], which divide SQA methods into "rule-based", "machine learning-based", and "deep learning-based", are subsets of Lucafo's categorization. According to this classification, rule-based methods computes several features from PPG, then use a set of thresholds to distinguish corrupted signals from reliable ones.

One example of this approach was proposed by [Vadrevu and Manikandan 2019] that introduced SQC by computing the absolute amplitude, zero-crossing rate, and autocorrelation features. [Reddy et al. 2020] presented another rule-based technique to reduce false alarms through the detection of signal saturation and sensor saturation. Differently to rule-based methods, the methods based on ML and Deep Learning (DL) extracted various features and then fed these features into a classifier to distinguish PPG segments into quality classes. For instance, [Chong et al. 2014] employed a support vector machine using time-domain features to perform the signal classification. Similarly, [Pereira et al. 2019] use support vector machine, k-nearest neighbors, and decision trees to model a SQC. [Azar et al. 2021] proposed the use of a deep autoencoder to perform PPG SQA.

From the literature, it is noticeable that learning-based methods are usually superior in terms of classification precision. However, these methods are commonly complex, computationally expensive, and time-consuming, in a way that in this case it is necessary to analyze the trade-off between computational cost, algorithm performance and storage availability on the wearable device. On the other hand, rule-based methods are fast and easy to implement, although their lack of generalization to unseen signal data. This generalization issue arises from the fact that their decision rules are formulated based on anticipated signal characteristics, and their thresholds are determined through empirical methods. In other words, these rule-based models have been crafted according to specific data characteristics, leading to reduced accuracy when confronted with unseen data.

In order to take advantage of the best aspects of both approaches and mitigate the drawbacks of each, [Lucafo et al. 2022] proposed a hybrid SQA model based on both rules and learning. Lucafo's method aims at reducing power consumption, which is highly desirable for constrained hardware, using a CNN only in specific cases when the quality of the signal is not easily determined by its waveform. In spite of their effort, Lucafo's method presented negligible influence on the quality metrics, even though reducing the estimated energy consumption for PPG SQA. Taking this into account, [Lima et al. 2023] contemplated a novel approach based on Neural Architecture Search (NAS) to find a DL model that consumes few computational resources and does not require specific thresholds or previously known signal characteristics to classify signals. Despite the similarity in approach, our goal is to discover a model for performing SQA, whereas [Lima et al. 2023] employed NAS to discover a neural architecture for classifying whether a given peak in PPG signals is systolic or not.

Finally, in this paper, we propose and explore the use of attention-based approaches [Bahdanau et al. 2015, Luong et al. 2015, Vaswani et al. 2017] in conjunction with the aforementioned NAS strategy to find an efficient SQC under strict memory-performance trade-off and power consumption considerations, enabling us to develop a compact SQC for deployment in embedded devices and real-time HMA.

2. Proposed Method

PPG offers numerous advantages to wearable-based health monitoring applications, especially those worn on the wrist. It is easy to configure, convenient, low in complexity, and cost-effective. Modern PPG devices commonly feature a solitary optical sensor that combines a near-infrared emitter and detector. This sensor is incorporated into a reusable and comfortable design that seamlessly integrates with computational resources. This

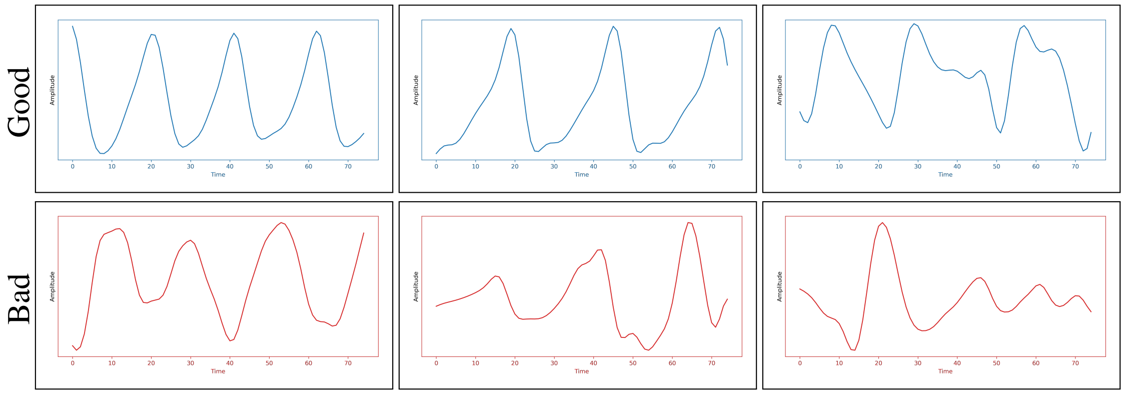


Figure 1. Illustrations of reliable and unreliable PPG signals. A 'reliable' signal exhibits symmetric, well-defined, and periodic patterns. Conversely, 'unreliable' signals shows asymmetric patterns with reduced redundancy and increased entropy between intervals. These visualizations were generated using signals from the ICON dataset ([Fioravanti et al. 2024]).

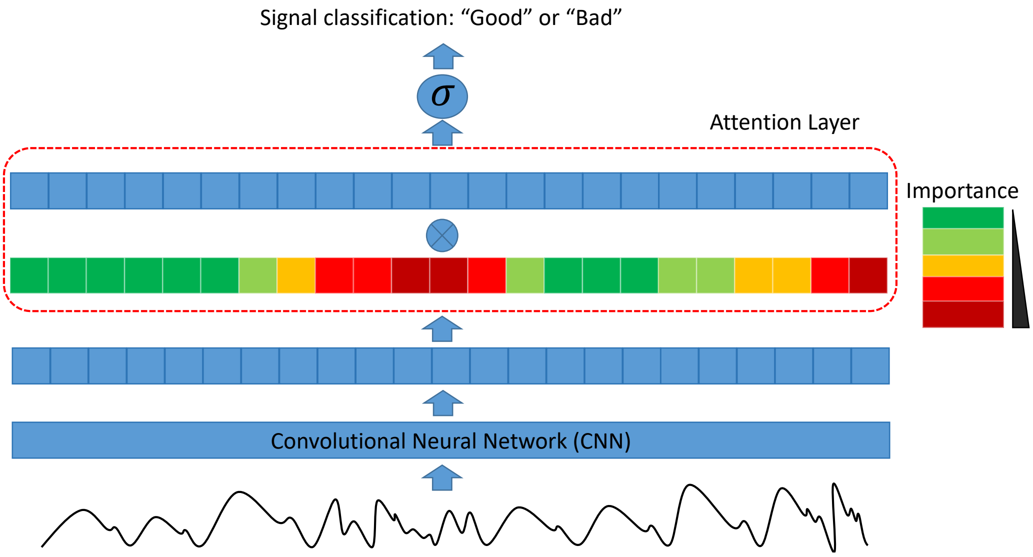


Figure 2. Our attentive-CNN method for PPG signal quality classification. We combine an optimized convolutional model with additional attention mechanism steps to improve the ability of the resulting model in distinguishing high quality segments from low quality ones, so that only the high quality ones may be used for extracting reliable indicators, such as heart rate, in real-time wearable applications.

configuration enables uninterrupted measurements, making it beneficial for continuous monitoring solutions in wearable electronics, including fitness trackers and smartwatches. Despite of these advantages, PPG signals are susceptible to significant degradation caused by various factors, particularly motion artifacts. Excessive movement of the PPG sensor may cause morphological disruptions in the waveform, thereby affecting the subsequent analysis of the signal. Figure 1 shows some examples of reliable (with significant information) and unreliable (disrupted) PPG signals. Unreliable signals can lead to erroneous decision-making and misjudgments, which is unacceptable for HMA. Hence, methods for assessing the quality of PPG signals are crucial to avoid misinterpretation, enabling the differentiation between reliable and noisy signals. Our work builds on ML models to generate robust and well-performing models to classify PPG segments for SQA estimation purposes. We employ NAS techniques, combined with attention mechanisms, to discover neural network solutions that are lightweight yet deliver competitive performance within strict memory and computational limitations, as succinctly depicted in Figure 2.

CNN were initially proposed for Computer Vision domain, as a way of automatically extracting local and global features of a given image or frame for posterior tasks, such as classification. Their usage eventually extended to other domains, such as time series classification, including biomedical ones [Lucafo et al. 2022]. On the other hand, attention mechanisms were initially proposed in the Natural Language Processing domain, as a way of circumventing the commonly found problem of vanishing gradients in long-term dependency learning of sequential models[Bahdanau et al. 2015], such as Recurrent Neural Networks, and found widespread applicability, as well as state-of-the-art performance, in several domains, such as Large Language Models[Brown et al. 2020] and Vision-related tasks[Freitas et al. 2023a].

Our goal is to improve the capabilities of a compact CNN with the assistance of attention mechanisms for quality classification of PPG sensor data, which is a functionality widely employed in several wearable applications, pertaining to heart rate and sleep assessment, just to mention a few. Our baseline CNN, described in Figure 3a, was discovered to a NAS procedure similar to the one in [Lima et al. 2023], albeit focused in the PPG signal quality assessment, and not on systolic peak classification. Our three proposed additional attention mechanisms, namely, Additive Attention Layer, Dot Product Attention Layer and Scaled Dot Product Attention Layer, will be further described in the following subsections.

2.1. Additive Attention Layer

The first proposed attention mechanism, namely Additive Attention Layer (AAL), was introduced by [Bahdanau et al. 2015] as a way to compute context vectors for captured relevant information from a sequential input. This mechanism uses a linear transformation followed by a nonlinear activation function:

$$\Phi(\boldsymbol{s}) = \psi(\boldsymbol{W} \odot \boldsymbol{s} + \boldsymbol{b}), \tag{1}$$

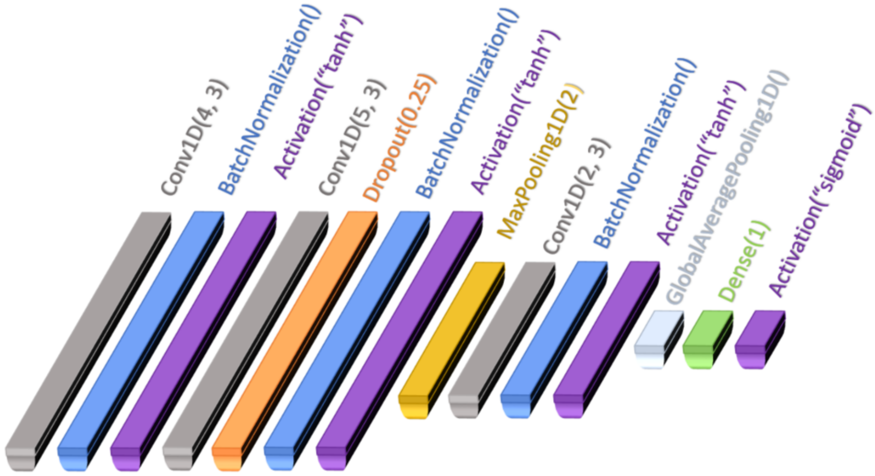
where $\Phi(\cdot)$ is denoted the "Bahdanau Function", $\boldsymbol{W}$ and $\boldsymbol{b}$ are learnable vectors of the same dimensions as the $\boldsymbol{s}$ vector, which corresponds to the output of the previous layer, and $\odot$ represents the element-wise product. $\psi(\cdot)$ is a nonlinear activation function, here chosen as hyperbolic tangent. The "Bahdanau Function" output is further transformed by a "Softmax function", to obtain the *Attention Scores* $\boldsymbol{\eta}(\boldsymbol{s})$:

$$\boldsymbol{\eta}(\boldsymbol{s})_i = \sigma(\Phi(\boldsymbol{s}))_i = \frac{\mathrm{e}^{\Phi(\boldsymbol{s})_i}}{\sum_j \mathrm{e}^{\Phi(\boldsymbol{s})_j}}, \tag{2}$$

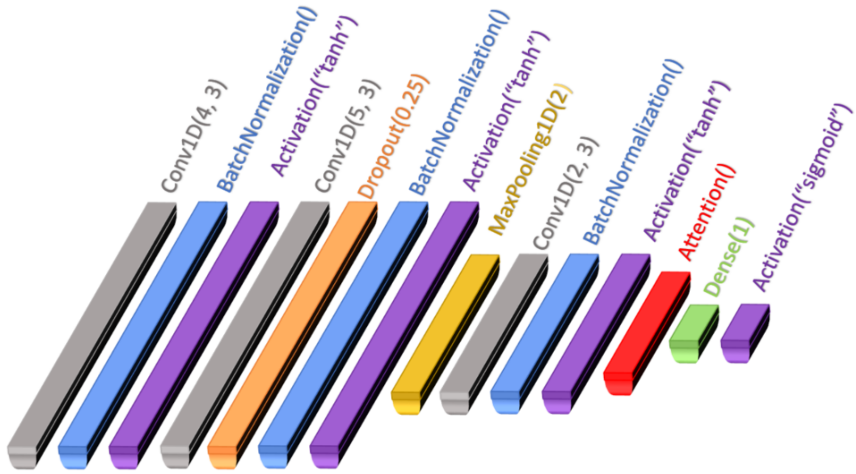
Finally, we multiply the same encoding vector $\boldsymbol{s}$ by the attention scores, so that we can obtain the *context vectors* $\boldsymbol{z}$:

$$\boldsymbol{z} = \boldsymbol{s} \odot \boldsymbol{\eta}(\boldsymbol{s}) \tag{3}$$

Our work consists of replacing the Global Pooling Layer (GPL) in the original model (see Figures 2 and 3b) by the AAL.



(a) The baseline CNN model discovered via Neural Architecture Search using the ICON dataset ([Fioravanti et al. 2024]). The network processes information sequentially through the layers without incorporating an attention mechanism.



(b) Our proposed improved attention-based model. The attention mechanism enables the network to focus on specific parts of the input data, enhancing its ability to learn and process information more effectively.

Figure 3. Comparison of Neural Network Architectures with and without Attention Mechanism

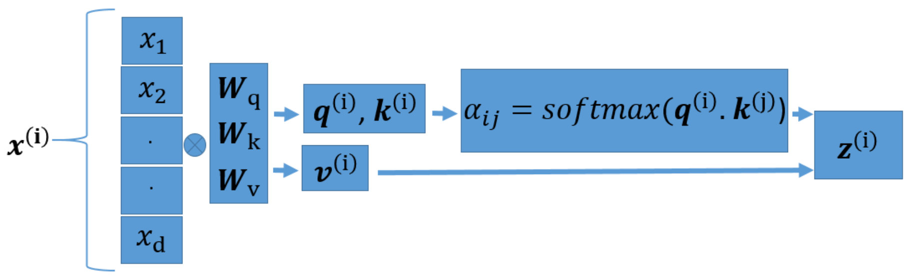


Figure 4. Dot Product Attention Layer (DPAL). The softmax function in uses the result of the dot product between the query and key vectors as argument.

2.2. Dot Product Attention Layer

[Vaswani et al. 2017] proposed DPAL, a crucial component in many deep learning models, particularly in Natural Language Processing (NLP) tasks. DPAL computes attention scores between a query vector and a set of key vectors, typically derived from the input sequence. These attention scores represent the importance or relevance of each element in the input sequence to the query. It works with basis on four concepts:

- Each token in the input sequence is associated with three vectors: a query vector, a key vector, and a value vector. These vectors are typically learned during training.
- *Dot product:* For each query vector, the dot product is computed with every key vector. This results in a set of scores representing the similarity between the query and each key.
- *Softmax*: The dot product scores are inputs of softmax function to obtain a distribution of attention weights, ensuring that they sum up to 1 and represent the importance of each key vector relative to the query.
- *Weighted Sum*: Finally, the attention weights are used to compute a weighted sum of the corresponding value vectors. This weighted sum represents the representation of the input sequence with respect to the query.

Mathematically, DPAL can be expressed as $attention(Q, K, V)$=$softmax\ (QK^{\intercal})\ V$, where Q represents the query matrix, K represents the key matrix, and V represents the value matrix. The DPAL allows the model to efficiently capture long-range dependencies in the input sequence by attending to relevant information, making it a powerful tool for various NLP tasks.

2.3. Scaled Dot Product Attention Layer

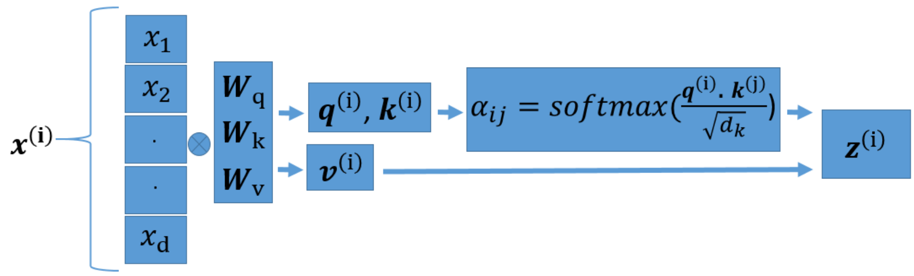


Figure 5. Scaled Dot Product Attention Layer (SDPAL). The softmax function uses the result of the dot product between the query and key vectors multiply by a normalization parameter d_k as argument.

The SDPAL is a variation of DPAL which apply a normalization factor $\sqrt{\boldsymbol{d}_k}$ (see Figure 4), and was also described in [Vaswani et al. 2017]. As in the case of AAL, we also replaced the GPL of Figure 3a with a distinct operation, the SDPAL, which is slightly more complex, and requires three additional learnable vectors: *query* ($\boldsymbol{q}$), *key* ($\boldsymbol{k}$) , with the query and the key vectors of the same dimensions, given as a hyperparameter of the model, and vector *values* ($\boldsymbol{v}$), with a size which is also given as a hyperparameter. Its steps are concisely described in Figure 5. As in the case of AAL, these auxiliary variables aim to map the cross-importance between each element of the sequence input, generating as output the i^{th} attention weight of each corresponding j^{th} element of the sequence data. The $\boldsymbol{d}_k$ is a normalization operation over the output from the previous steps.

3. Experimental Setup

We evaluate the performance of the proposed architectures described in Figure 3 by comparing the found neural network classifiers with the state-of-the-art SQC using the same procedure. This comparison was conducted using the ICON dataset [Fioravanti et al. 2024]. This dataset comprises 46 subjects, corresponding to PPG signals collected at 25Hz, lasting 45-60 minutes per subject using a *Samsung Smartwatch Galaxy Watch Active 2*©. These subjects include 9 volunteers with permanent atrial fibrillation, 16 volunteers with normal sinus rhythm, and 21 volunteers with other non-specified arrhythmias. For each subject, the PPG sensor on the non-dominant wrist was strategically placed over the radial artery to ensure optimal signal capture. The data collection adhered to a protocol under resting conditions, with participants advised to either sit or recline in a calm environment. The PPG signals, capturing baseline heart rate and blood flow, were recorded during periods of relaxation. The quality of the signals in the ICON dataset was manually labeled by experts according to their waveforms and according to the aim of this dataset, which is to allow the measurement of Interbeat Intervals (IBIs).

The PPG signals were divided into 3-second windows, each consisting of 75 samples with a 5-sample overlap. These segments of signal were normalized using min-max normalization within the range $[0, 1]$. In preparation for the learning phase, each segment was labeled based on a threshold d representing the proportion of samples annotated as high quality within the segment. Subsequently, the segment was categorized as 'reliable' if the fraction of high-quality signal exceeded d, and 'unreliable' otherwise. Specifically, $d = 0.8$ was used to create these labels in the experiments.

In our experiments, we split the data into training and test sets. The validation subsets were randomly sampled as a subset of training set. During the training phase, we utilize the data and corresponding annotations from 64% of the subjects within each assessment group. For validation purposes, we opt for data from 16% of the subjects in each assessment group. The remaining 20% of the subjects in each group are employed to test the proposed methods and the overall pipeline. It is important to emphasize that the signal segments comprising the experiments reported here are all grouped according to the subject to which the signal belongs. This means that signal segments from subjects in the training set do not overlap with those in the testing set, and vice versa. This subject-based split was chosen to prevent training bias from also being present in testing. This type of split allows us to simulate how the model would perform with an individual who is not present in the dataset.

The DL models were implemented using Keras 2.8.0 and Python 3.8. All models were trained using 20 epochs. To test the performance of the proposed and the state-of-the-art methods, we compare the predicted quality indices with the pre-labeled indices provided in the benchmark dataset using the accuracy, precision, recall, and f1-score metrics. More details about these metrics are available at [Dalianis and Dalianis 2018].

4. Experimental Results

Table 1 shows the results obtained using the compared state-of-the-art methods, namely 'Lucafo$_1$', 'Lucafo$_2$', 'RP+ViT', 'MTF+ViT', and 'Hao & Bo'. All results in this table were attained using the ICON dataset with the same experimental conditions. From this table, the first noticeable achievement concerns the performance of the baseline. By

Table 1. Performance comparison of the baseline model (Figure 3a), its variations including an attention layer (Figure 3b), and state-of-the-art methods using ICON dataset.

Method	Reference	Accuracy(%)	Precision(%)	Recall(%)	F1-Score(%)
Lucafo_1	[Lucafo et al. 2022]	89.8	88.6	95.9	91.9
Lucafo_2	[Lucafo et al. 2022]	89.9	87.5	**97.1**	92.0
Hao & Bo	[Hao and Bo 2021]	80.6	75.7	90.5	82.4
RP+ViT	[Freitas et al. 2023b]	89.9	95.1	90.3	92.6
MTF+ViT	[Freitas et al. 2023b]	90.3	91.2	94.1	92.6
NAS-CNN (Baseline)	[Lima et al. 2023]	92.2	95.1	93.0	94.1
NAS-CNN + AAL		**92.7**	95.1	93.9	**94.5**
NAS-CNN + DPAL		91.7	**95.6**	91.9	93.7
NAS-CNN + SDPAL		92.1	95.4	92.5	94.0

adopting the approach described by [Lima et al. 2023] but adapting it to our problem, we discovered a tiny CNN that is able to surpass the state-of-the-art in almost all performance metrics but notably in terms of accuracy, precision, and f1-score. Moreover, based the results, we can also notice the benefit of the addition of attention layers. In other words, the addition of an attention layer to the baseline model effectively improved all its performance metrics. Specifically, the AAL exhibited significant improvements with only a minimal increase in the total number of model parameters. While SDPAL and DPAL also delivered competitive results, they necessitated a higher number of parameters.

5. Conclusions

In this paper, we investigated an approach for generating efficient CNN architectures to classify whether a given PPG signal segment is suitable for use or not using the ICON dataset. This task is addressed as a binary classification problem and is crucial to enable reliable HMA, especially in the context of wearables devices. Using NAS-based technique, we found a baseline CNN architecture that are tiny enough to be deployed in embedded devices to enable real-time wearable HMA. Moreover, we show that the used baseline CNN had its performance improved through additional attention layers, especially for the AAL, which proved to be efficient in its proposal to improve the baseline accuracy without increasing the computational cost of the application.

References

Azar, J., Makhoul, A., Couturier, R., and Demerjian, J. (2021). Deep recurrent neural network-based autoencoder for photoplethysmogram artifacts filtering. *Computers & Electrical Engineering*, 92:107065.

Bahdanau, D., Cho, K., and Bengio, Y. (2015). Neural machine translation by jointly learning to align and translate. *3rd International Conference on Learning Representations, ICLR 2015*.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I., and Amodei, D. (2020). Language models are few-shot learners. In Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M., and Lin, H., editors, *Advances in Neural Information Processing Systems*, volume 33, pages 1877–1901. Curran Associates, Inc.

Chatterjee, T., Ghosh, A., and Sarkar, S. (2022). Signal quality assessment of photoplethysmogram signals using quantum pattern recognition technique and lightweight cnn module. In *International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 3382–3386.

Chong, J. W., Dao, D. K., Salehizadeh, S., McManus, D. D., Darling, C. E., Chon, K. H., and Mendelson, Y. (2014). Photoplethysmograph signal reconstruction based on a novel hybrid motion artifact detection–reduction approach. part i: Motion and noise artifact detection. *Annals of biomedical engineering*, 42:2238–2250.

Cisco (2019). Cisco visual networking index (vni) global mobile data traffic forecast update, 2017-2022 white paper. White paper. Accessed on 15 Mar 2021.

Dalianis, H. and Dalianis, H. (2018). Evaluation metrics and evaluation. *Clinical Text Mining: secondary use of electronic patient records*, pages 45–53.

Elgendi, M. (2016). Optimal signal quality index for photoplethysmogram signals. *Bioengineering*, 3(4):21.

Fioravanti, V. B. O., Freitas, P. G., Rodrigues, P. G., de Lima, R. G., Lucafo, G. D., Cabello, F. C., Seidel, I., and Penatti, O. A. B. (2024). Machine learning framework for inter-beat interval estimation using wearable photoplethysmography sensors. *Biomedical Signal Processing and Control*, 88:105689.

Freitas, P. G., De Lima, R. G., Lucafo, G. D., and Penatti, O. A. (2023a). Assessing the quality of photoplethysmograms via gramian angular fields and vision transformer. In *31st European Signal Processing Conference (EUSIPCO)*, pages 1035–1039.

Freitas, P. G., Lima, R. G. D., Lucafo, G. D., and Penatti, O. A. B. (2023b). Photoplethysmogram signal quality assessment via 1d-to-2d projections and vision transformers. In *International Conference on Quality of Multimedia Experience, Ghent, Belgium, June 20-22*, pages 165–170.

Gambarotta, N., Aletti, F., Baselli, G., and Ferrario, M. (2016). A review of methods for the signal quality assessment to improve reliability of heart rate and blood pressures derived parameters. *Medical & biological engineering & computing*, 54:1025–1035.

Gartner (2021). Gartner forecasts global spending on wearable devices to total $81.5 billion in 2021. Accessed on 15 Mar 2021.

Hao, J. and Bo, G. (2021). A quality assessment system for ppg waveform. In *International Conference on Circuits and Systems (ICCS)*, pages 170–175.

Li, K., Warren, S., and Natarajan, B. (2011). Onboard tagging for real-time quality assessment of photoplethysmograms acquired by a wireless reflectance pulse oximeter. *IEEE Transactions on Biomedical Circuits and Systems*, 6(1):54–63.

Li, Q. and Clifford, G. D. (2012). Dynamic time warping and machine learning for signal quality assessment of pulsatile signals. *Physiological measurement*, 33(9):1491.

Lima, R. G., Freitas, P. G., Lucafo, G. D., Fioravanti, V., Seidel, I., and Penatti, O. A. (2023). Neural architecture search for tiny detectors of inter-beat intervals. In *European Signal Processing Conference (EUSIPCO)*, pages 1085–1089.

Lucafo, G., Freitas, P. G., Lima, R., Luz, G., Bispo, R., Rodrigues, P., Cabello, F., and Penatti, O. (2022). Signal quality assessment of photoplethysmogram signals using hybrid rule- and learning-based models. In *XIX Congresso Brasileiro de Informática em Saúde (CBIS)*. Brazilian Health Informatics Association (SBIS).

Luong, T., Pham, H., and Manning, C. D. (2015). Effective approaches to attention-based neural machine translation. *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 1412–1421.

Naeini, E. K., Azimi, I., Rahmani, A. M., Liljeberg, P., and Dutt, N. (2019). A real-time ppg quality assessment approach for healthcare internet-of-things. *Procedia Computer Science*, 151:551–558.

Pereira, T., Gadhoumi, K., Ma, M., Liu, X., Xiao, R., Colorado, R. A., Keenan, K. J., Meisel, K., and Hu, X. (2019). A supervised approach to robust photoplethysmography quality assessment. *IEEE journal of biomedical and health informatics*, 24(3):649–657.

Reddy, G. N. K., Manikandan, M. S., and Murty, N. N. (2020). On-device integrated ppg quality assessment and sensor disconnection/saturation detection system for iot health monitoring. *IEEE Transactions on Instrumentation and Measurement*, 69(9):6351–6361.

Selvaraj, N., Mendelson, Y., Shelley, K. H., Silverman, D. G., and Chon, K. H. (2011). Statistical approach for the detection of motion/noise artifacts in photoplethysmogram. In *International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4972–4975.

Sukor, J. A., Redmond, S., and Lovell, N. (2011). Signal quality measures for pulse oximetry through waveform morphology analysis. *Physiological measurement*, 32(3):369.

Sun, X., Yang, P., and Zhang, Y.-T. (2012). Assessment of photoplethysmogram signal quality using morphology integrated with temporal information approach. In *International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 3456–3459.

Vadrevu, S. and Manikandan, M. S. (2019). Real-time ppg signal quality assessment system for improving battery life and false alarms. *IEEE transactions on circuits and systems II: express briefs*, 66(11):1910–1914.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems*, pages 5998–6008.

Zhang, X., Li, J., Cai, Z., Zhao, L., and Liu, C. (2022). Deep learning-based signal quality assessment for wearable ecgs. *IEEE Instrumentation & Measurement Magazine*, 25(5):41–52.