Don't Sweat It: Transforming Electrodermal Activity for Stress Detection

Mjellma Çitaku, Jae-Eun Nam, Larissa Zott, Silja Meyer-Nieberg
Institute for Techincal Informatics, University of the Bundeswehr Munich,
Munich, Germany
mjellma.citaku@unibw.de, jae-eun.nam@unibw.de, larissa.zott@unibw.de, silja.meyer-nieberg@unibw.de

***Abstract*—With stress being a commonly occurring emotion in society as well as a major component in causing serious physical and mental health issues, studying and monitoring this phenomenon is an important task in the modern world. Recent studies have shown the success of using deep learning approaches in stress classification. A promising way to proceed is to transform the one-dimensional biophysiological time series into two-dimensional representations by image encoding. However, the question arises which encoding type is suitable for which kind of time series. Here, several often used image encodings are analyzed in regards to their ability to extract and display signal characteristics relevant for stress detection. In this paper, the focus lies on the electrodermal activity (EDA) which represents a promising modality to detect stress. Due to data scarcity, the utility of transfer learning starting from ImageNet-pretrained convolutional neural networks (CNNs) is also analyzed. Here, several backbone families, established and modern, are evaluated under a locked fine-tuning protocol to isolate architectural effects. Furthermore, the impact of tonic-phasic decomposition is examined by comparing the full EDA signal to its phasic component. Experiments on the WESAD dataset demonstrate that scalograms appear more robust for stress detection. Additionally, the trade-off between lightweight and larger pretrained CNNs is characterized.**

Index terms **stress detection, CNN, image encodings, deep learning, frugal, electrodermal activity, WESAD**

I. Introduction and Related Work

Stress can have a severe harmful impact, be it short-term or long-term in its chronic form [1]. On the one hand, a person's quality of action and decision making may be affected by stress momentarily. On the other hand, stress in its chronic form may lead to serious decline of mental [2], [3] or physical health [4], [5]. The detection of stress is therefore highly important. This work explores the detection of stress based on biophysiological signals - specifically the electrodermal activity (EDA). Certain biosignals, among them EDA, have shown a correlation to a human's stress state [6], [7], while their response to stimuli can significantly differ between individuals. This can be attributed to factors such as age [8] and gender [8], [9].

This paper is funded by dtec.bw – Digitalization and Technology Research Center of the Bundeswehr and the project KOIOS [101103770]. dtec.bw is funded by the European Union – NextGenerationEU, and KOIOS is a project funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Defence Fund. Neither the European Union nor the granting authority can be held responsible for them.

In this paper, convolutional neural networks are used to determine a subject's stress state using EDA time series signals. The EDA signal is often utilized for the detection of stress since the sweat glands are influenced by the sympathetic nervous system [10]. It can be further decomposed into a phasic and a tonic component. The phasic part or skin conductance response varies fast in response to stimuli whereas the slowly changing tonic part or skin conductance level is the baseline conductivity component [11].

Commonly, either machine learning (ML) or deep learning (DL) techniques are used to distinguish between stress and non stress states. In the case of the former, hand-crafted features as median or maximum peak value serve as input for training ML systems, e.g. [12], [13]. In the case of DL variants, a multitude of different techniques have been employed, see e.g. [14]. One promising technique is to encode the time series as an image, called image encoding (IE), thus enhancing certain characteristics of the signal [15]. Often, feature extraction is conducted with convolutional neural networks (CNN) [16], [17], [18]. Several encodings have been investigated. For example, [19] considered recurrence plots whereas [16] focused on continuous wavelet transforms (CWT) and the resulting scalograms for the detection of stress in drivers. [17] combined several IE variants that preserve different characteristics of the time series as well as different modalities while [18] investigated unified images of three modalities.

This paper presents an analysis of CNN architectures and transfer learning for image encodings of physiological time series: Different pretrained DL architectures were fine-tuned using five different IEs of the EDA signal of multiple subjects as input. Since each encoding variant extracts different characteristics of the signal, their performance as CNN input was compared to assess how suitable the respective characteristics are in terms of the EDA signal. Before applying the image transformation, the signals were first divided into segments of equal length. Furthermore adhering to the current state-of-the art, leave-one-subject-out cross-validation (LOSO) was applied.

The work is based on the findings in [20], where several CNN architectures were analyzed for the blood volume pulse (BVP), another modality often used in stress detection. There, scalograms and an EfficientNet variant achieved the best results. Here, the analysis is extended by focusing on EDA which in contrast to the BVP represents a non-periodic signal leading to the question whether a different representation of the time series is preferable. Furthermore, it considers different CNN architectures and architecture sizes. All in all, this paper addresses the following research-questions:

- Are certain CNN architecture design choices more suitable for the specific nature of time series-based IEs which differ strongly from natural images?
- Does increasing CNN model capacity and ImageNet pretraining improve stress detection for EDA IEs, or are lightweight architectures sufficient?
- Which image-based representations of EDA time series capture stress-related patterns for CNN-based classification best?

The paper is structured as follows: first, the details of the approach and the experimental setup are provided in Section II. The results and findings of the experiments are then presented and discussed in Section III. Section IV presents a summary of findings of this paper and points for future work.

II. Methodology

A. Data

The analysis considers the publicly available WESAD dataset by Schmidt et al. [21], which has been commonly used in stress detection studies. It comprises biophysiological and motion data from 15 people recorded by two different devices during specific lab study protocols. These protocols included tasks that were meant to induce emotion conditions such as stress or amusement, as well as a baseline recording. The investigation focuses on the EDA data gathered by the RespiBAN chest belt at a sampling frequency of 700 Hz in a non-invasive manner. We note that this is not the frequency that was finally used in our work.

B. Two-Dimensional Representations of Time Series: Image Encodings

As stated previously, image encodings, that is, two-dimensional representations of the EDA time series are considered. To derive the encodings, each class, baseline, amusement, and stress, was treated separately. Afterwards, the signals were divided into segments of identical length which were then transformed into image encodings. This was carried out for every encoding type considered. Each variant extracts specific information from the time series and displays it in a two dimensional image leading to distinct looks and patterns. The following IEs were implemented in this work: scalograms [22] (pp. 1-9, 313-325), Gramian angular summation and difference field (GASF and GADF respectively [15]), Markov transition field (MTF) and recurrence plots (RP) [23]. This selection represents the image encodings that are most often used for physiological time series.

Scalograms contain frequency information of the time series signal as they visualize the output value of the CWT. It can be seen as a map of the magnitude of the respective wavelet function's values as a function of time [22] (pp. 1-9, 313-325). In a GAF [15] transformation the time series is first translated into polar coordinates, where the time stamp is represented by the radius coordinate and the signal-value respectively by the angle coordinate. This will subsequently help visualize the time dependency between data points. The so-called Gramian matrix will then contain the cosine of the sum (summation field) or difference (difference field) of the angles of each point, therefore capturing the temporal correlation. The concept of GAFs was introduced by [15], as well as the Markov transition field. A MTF reflects the Markov transition probabilities between discrete bins in the time series. These probabilities are represented in a Markov transition matrix, by using a Markov chain operation of the first order along the time-axis. Due to the sequential nature of this process, the signal's time dependency is reflected in the final image [15].

Recurrence plots are a visualization of a system's recurrence [23] - a concept introduced by Henri Poincaré [24] in 1890. The latter describes the approximate return of a dynamic system to a previous state in phase space over time, regardless of its linearity or dimensionality. Hence, an RP visually reflects the recurrence of the signal investigated and its deterministic structure. The main diagonal in each image is called "line of identity" [23]. Fig. 1 shows a schematic representation of how a time series segment is displayed as an IE.

TABLE I
BACKBONE FAMILIES USED IN THIS WORK WITH RESPECTIVE PARAMETER COUNTS (IMAGENET-PRETRAINED).

Family	Backbone	# Params
MobileNetV3 [26]	MobileNetV3-Small	2.6M
	MobileNetV3-Large	5.5M
EfficientNetV2 [27]	EfficientNetV2-S	21.4M
	EfficientNetV2-M	54.4M
	EfficientNetV2-L	119M
ResNetV2 [28]	ResNet50V2	25.6M
	ResNet101V2	44.7M
ConvNeXt [29]	ConvNeXt-Tiny	28.6M
	ConvNeXt-Small	50M
	ConvNeXt-Base	88.6M

C. Experimental Setup

Backbone selection: Representative CNN backbones were selected from multiple architecture families (MobileNetV3, EfficientNetV2, ResNetV2, and ConvNeXt) to study whether large pretrained CNNs provide a measurable benefit for EDA-based stress classification. In each family, backbones of different sizes were included, ranging from lightweight to large variants (Table I), enabling comparisons across both architectural design choices and parameter scale. The same fine-tuning protocol was used across all backbones. Furthermore, they were all instantiated using the official Keras model implementations (Keras Applications) [25].

Preprocessing of data: The RespiBAN EDA signal from the WESAD dataset was used for all experiments. The raw EDA time series was denoised using a 6th-order Butterworth low-pass filter with a 1 Hz cutoff frequency [30]. The filtered

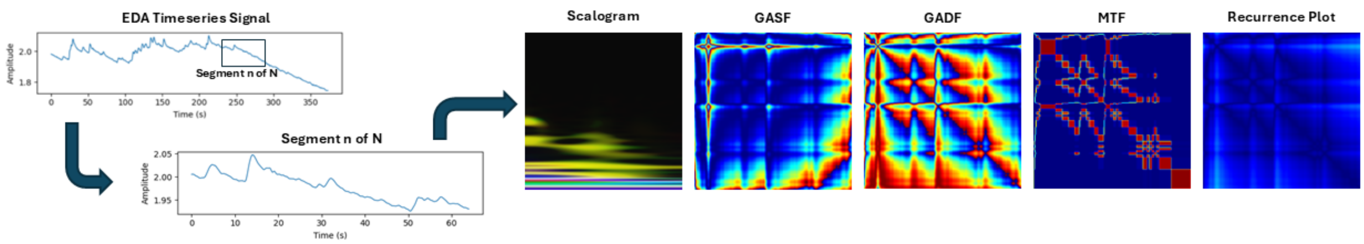

Fig. 1. Symbolic representation of an EDA signal being segmented and transformed into different 2D images.

signal, originally sampled at 700 Hz, was then resampled to 16 Hz to reduce compute while preserving the EDA dynamics [31]. Depending on the experimental condition, either the resampled EDA signal itself or its phasic component (EDA_phasic) obtained via the tonic-phasic decomposition [32] is used. Per-subject z-score standardization was applied using the session-wide mean and standard deviation computed prior to windowing. The standardized time series was segmented into 60 s windows with 0.75 overlap, and each window was converted into the IEs considered in this work. Scalograms used a CWT (python package: ssqueezepy [33]) with the generalized Morse mother wavelet (gmw), 16 voices per octave and JET-colored. Gramian angular fields were produced with the python package pyts [34] defaults in both GADF and GASF modes. For the Markov transition field pyts with 8 quantization bins was used. For the recurrence plots we employed the standard pyts configuration. All images were generated per window at 299×299 resolution. After segmentation and encoding, 1120 baseline, 610 stress, and 316 amusement images were obtained (2046 total). Unless stated otherwise, this preprocessing, windowing, and encoding pipeline was kept identical across all experiments.

Evaluation protocol and metrics: Subject-independent performance was evaluated using the standard LOSO cross-validation across the 15 subjects. In each fold, one subject was held out for testing, while the remaining subjects were used for training, with 20% of the training data reserved for validation. To capture training stochasticity, each experiment was repeated with 20 seeds, and the same set of seeds was used for all experiments. For a fixed seed, the LOSO results were first aggregated across folds by taking the median score over the 15 held-out test subjects. The mean $\pm$ standard deviation of the accuracy and weighted F1-score metrics was then computed across the seeds. To further assess whether the observed differences in performance were statistically meaningful, we also conducted statistical testing. We aggregated each model's 20 runs per subject to a median score (weighted F1 and accuracy). On these subject-wise medians we ran a Friedman test (α=0.05) with a post-hoc Nemenyi all-vs-all test procedure, reporting average ranks and the single critical difference (CD). The rank gaps above CD are marked significant.

Fine-tuning procedure: A lightweight classification head (Dropout $\rightarrow$ Dense $\rightarrow$ ELU $\rightarrow$ softmax) was attached to each pretrained backbone. Training was performed in two stages:

TABLE II
HYPERPARAMETER SEARCH SPACE FOR CLASSIFIER-HEAD TRAINING (CH) AND FINE-TUNING (FT). ES STANDS FOR EARLY STOPPING AND LR FOR LEARNING RATE.

Hyperparameter	**CH**	**FT**
Epochs	$\{10, 30\}$	$\{30, 50\}$
ES Patience	$\{3, 7, 5\}$	$\{3, 7, 5\}$
Adam LR	$\{1e{-}4, 1e{-}3, 3e{-}3\}$	$\{1e{-}5, 3e{-}5, 1e{-}4\}$

(i) classifier-head training with the backbone frozen (CH), followed by (ii) end-to-end fine-tuning with all layers unfrozen (FT). The Adam optimizer was used in both stages.

Model selection and hyperparameter tuning: Before running the full LOSO evaluation, the training hyperparameters were selected using a computationally efficient *mini-LOSO* protocol, where subjects S7 and S14 acted as held-out test subjects. Mini-LOSO experiments were repeated with 15 seeds and used ResNet50 [35] as a fixed reference architecture. The number of epochs, the early stopping (ES) patience as well as the Adam learning rate were tuned for both fine-tuning stages (CH, FT) (Table II). Candidate configurations included combinations formed from the first and last values in each list, complemented by randomly sampled intermediate settings, resulting in 134 total configurations. The final configuration was selected by ranking candidates based on the weighted F1-score and the accuracy, and was then fixed for all subsequent experiments.

III. Results and Discussion

Based on the hyperparameter tuning described in Section II-C the configuration *CH: lr=1e-3, ES patience=5, epochs=10; FT: lr=1e-5, ES patience=7, epochs=30* yielded the best results in regards to the weighted F1-score and accuracy. The selected configuration was kept fixed for all backbones to ensure a fair comparison, such that observed differences can be attributed to the backbone choice rather than to unequal tuning effort or backbone-specific hyperparameters. In the following, results are reported for the 3-class WESAD classification under LOSO across the considered backbone families, with accuracy and weighted F1-score summarized as mean $\pm$ standard deviation over 20 seeds (after median aggregation across folds).

The best-performing traditional ML method (linear discriminant analysis) from the original WESAD study [21] was used as a baseline. Using the manually extracted statistical

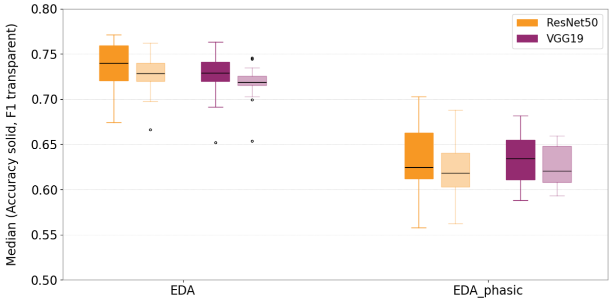

Fig. 2. Comparison of modeling EDA and its phasic component by accuracy (solid) and weighted F1-score (transparent) using ResNet50 and VGG19 as backbone architecture, respectively. For exact numbers, see Table III.

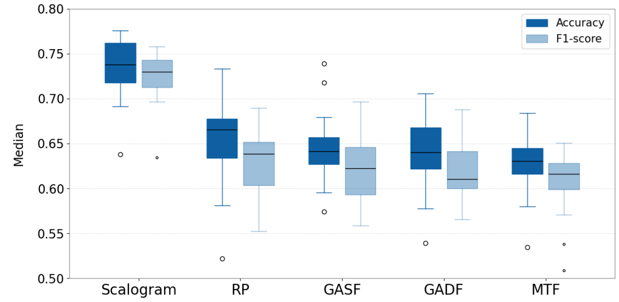

Fig. 3. Comparison of IEs (scalogram, recurrence plot (RP), Gramian angular summation/difference field (GASF/GAFM), Markov transition field (MTF)) by accuracy (solid) and weighted F1-score (transparent) using ResNet50V2. For exact numbers, see Table IV.

TABLE III
COMPARISON OF MODELING EDA AND ITS PHASIC COMPONENT BY ACCURACY AND WEIGHTED F1-SCORE (MEAN ± STANDARD DEVIATION, 20 RUNS).

Architecture	Signal	Accuracy	Weighted F1
ResNet50 [35]	**EDA**	**0.7371 ± 0.028**	**0.7263 ± 0.022**
	EDA_phasic	0.6321 ± 0.034	0.6195 ± 0.032
VGG19 [37]	**EDA**	**0.7258 ± 0.026**	**0.7180 ± 0.019**
	EDA_phasic	0.6344 ± 0.027	0.6268 ± 0.022

TABLE IV
COMPARISON OF IMAGE ENCODINGS BY ACCURACY AND WEIGHTED F1-SCORE (MEAN ± STANDARD DEVIATION, 20 RUNS). AVERAGE RANKS ARE REPORTED FOR THE WEIGHTED F1 SCORE (α=0.05, CD=2.23).

Image Encoding	Accuracy	Weighted F1	Avg. Rank
Scalogram	**0.7350 ± 0.033**	**0.7236 ± 0.028**	1.73
RP	0.6517 ± 0.048	0.6293 ± 0.037	2.83
MTF	0.6267 ± 0.037	0.6072 ± 0.036	3.67
GASF	0.6450 ± 0.038	0.6202 ± 0.037	3.47
GADF	0.6392 ± 0.041	0.6184 ± 0.034	3.3

features of the RespiBan EDA signals and applying LOSO, the approach achieves an accuracy of 0.67.

A. The Electrodermal Activity and its Phasic Component

The EDA signal is commonly decomposed into a tonic and a phasic component. The tonic component reflects the slowly varying baseline level of skin conductance (skin conductance level), while the phasic component captures short-term skin conductance responses associated with rapid sympathetic arousal [31]. Because stress is often characterized by sudden physiological changes, the phasic component is frequently assumed contain more relevant information for stress detection [36]. The impact of modeling the full EDA signal versus its extracted phasic component (EDA_phasic) is evaluated in Table III and Fig. 2. Since the training hyperparameters are selected using ResNet50 [35] as the reference backbone, VGG19 [37], is included as an additional, architecturally distinct model to verify that the observed conclusions are not specific to the tuning backbone. For both models, using only the phasic component results in a clear performance degradation of approximately 10%, accompanied by increased variance across runs. These findings indicate that the information loss introduced by signal decomposition outweighs the potential benefits of isolating phasic activity as a form of feature engineering. Consequently, retaining the full EDA signal appears more effective for stress classification in the examined setting.

B. Comparing Image Encodings

The performance of different IE methods using ResNet50V2 is compared in Table 3 and visualized in Fig. 3. Among the evaluated methods, scalograms achieved the highest performance with a mean accuracy around 0.75 and weighted F1-score around 0.72 across 20 runs. RP followed by a clear margin with a mean accuracy around 0.65 and weighted F1-score around 0.63, however with a notably higher variance. The remaining encodings, MTF, GASF and GADF, showed comparable but lower performance. As demonstrated in the boxplots in Fig. 3, scalogram encodings yielded both higher median scores and stronger robustness. The temporal-frequency characteristics captured in scalograms seemed to be most effective for stress detection in our setting. After confirming overall differences with the Friedman test, the Nemenyi post-hoc test showed a significant difference in weighted F1-score for the pairs scalograms vs. RP and scalogram vs. GASF. In terms of accuracy, the test agrees on significant differences for scalogram vs. RP.

C. Architecture Comparison Across Backbone Families

The comparison of the ImageNet-pretrained CNN backbones is summarized in Table V. Hereby we focused on the EDA signal and its scalogram representations, as they demonstrated the strongest performance in previous experiments. Overall, the strongest results are obtained by the ResNetV2 and EfficientNetV2 families, while ConvNeXt variants consistently under-perform in this setting. This trend is also illustrated in Fig. 4 that shows the distribution of the accuracy and weighted F1 across the 20 runs. The ResNetV2 and EfficientNetV2 groups remain centered around 0.73-0.76 median accuracy, whereas ConvNeXt lies around 0.61-0.65.

Among all evaluated backbones, the highest mean accuracy is achieved by ResNet101V2 and EfficientNetV2-M, with

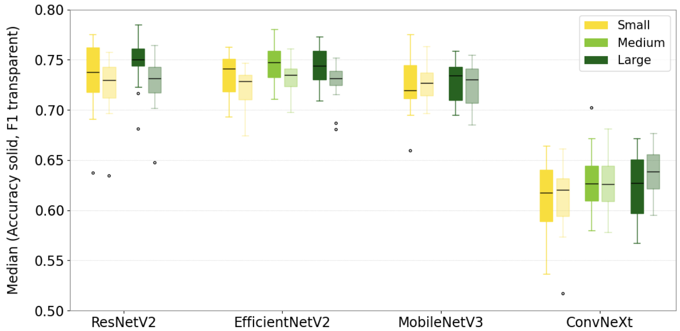

Fig. 4. Comparison of ResNetV2(50/101), EfficientNetV2(S/M/L), MobileNetV3(Small/Large), and ConvNeXt(Tiny/Small/Base) by accuracy (solid) and weighted F1-score (transparent). For exact numbers, see TABLE V.

TABLE V
COMPARISON OF ARCHITECTURES BY ACCURACY AND WEIGHTED F1-SCORE ON WESAD SCALOGRAMS (MEAN $\pm$ STANDARD DEVIATION, 20 RUNS). AVERAGE RANKS ARE REPORTED FOR THE WEIGHTED F1 SCORE ($\alpha = 0.05$, CD=3.33).

Architecture	Accuracy	Weighted F1	Avg. Rank
ResNet50V2 [28]	0.7350 $\pm$ 0.033	0.7236 $\pm$ 0.028	5.6
ResNet101V2 [28]	**0.7476** $\pm$ **0.022**	**0.7281** $\pm$ **0.026**	4.4
EfficientNetV2S [27]	0.7345 $\pm$ 0.021	0.7219 $\pm$ 0.020	5.07
EfficientNetV2M [27]	**0.7463** $\pm$ **0.019**	**0.7327** $\pm$ **0.016**	4.2
EfficientNetV2L [27]	**0.7432** $\pm$ **0.019**	**0.7279** $\pm$ **0.018**	3.8
MobileNetV3Small [26]	0.7265 $\pm$ 0.028	0.7281 $\pm$ 0.020	5.13
MobileNetV3Large [26]	0.7282 $\pm$ 0.019	0.7248 $\pm$ 0.020	5.2
ConvNeXtTiny [29]	0.6153 $\pm$ 0.034	0.6127 $\pm$ 0.032	7.13
ConvNeXtSmall [29]	0.6282 $\pm$ 0.030	0.6281 $\pm$ 0.026	6.73
ConvNeXtBase [29]	0.6248 $\pm$ 0.031	0.6382 $\pm$ 0.026	7.73

comparable weighted F1-scores. MobileNetV3 performs competitively, reaching 0.73 mean accuracy and 0.72-0.73 mean weighted F1. In contrast, all ConvNeXt variants achieve lower performance.

Within ResNetV2, the larger ResNet101V2 improves over ResNet50V2, indicating a modest benefit from increased capacity within this family. Within EfficientNetV2, the best overall performance is obtained by the mid-sized EfficientNetV2-M, while EfficientNetV2-L does not provide further gains and slightly underperforms V2-M. The Small variant remains competitive. Within MobileNetV3, only small differences are observed between the Small and Large variants. MobileNetV3-Large attains slightly higher accuracy, while MobileNetV3-Small yields comparable weighted F1. Within ConvNeXt, performance is low for all sizes, and scaling is not monotonic. ConvNeXt-Small achieves the highest accuracy among ConvNeXt variants, while ConvNeXt-Base yields the highest weighted F1. The results suggest that an increasing capacity does not necessarily lead to better results for this family under the current setup.

Across runs, the EfficientNetV2 family shows consistently low variability. ResNet101V2 is similarly stable, whereas ResNet50V2 is less consistent across the runs. For MobileNetV3, while the Small variant is more volatile, the Large variant is comparably stable to EfficientNetV2. The ConvNeXt variants display the largest scatter overall in addition to their lower mean performance. The Friedman test confirmed overall differences and the Nemenyi post-hoc analysis for the weighted F1-score identified significant differences when comparing ConNeXt-Base to EfficientNetV2L and EfficientNetV2M. Concerning the accuracy, ConvNeXt-Base differed significantly from EfficientNetV2L, ResNetV2101, and EfficientNetV2M.

IV. CONCLUSIONS AND OUTLOOK

This paper focused on stress detection based on physiological time series. Here, deep learning approaches have gained more and more importance. However, several challenges remain, among others, how to represent physiological time series and how to cope with the general data scarcity since available datasets are typically small. Image encodings or two-dimensional representations may enhance certain characteristics of the signal that allow to differentiate between states. Several image encodings have been introduced, however, no clear findings exist which variant should be preferred for which physiological time series. The paper focused on the electrodermal activity which is often used for stress detection. It evaluated several pretrained CNN architecture variants which were fine-tuned for the classification task. The EDA modality contains a phasic and tonic component with the former reacting swiftly to stressors. Using only the phasic part of the EDA signal, however, consistently degrades performance and yields less consistent outcomes across runs, indicating that the information removed by tonic-phasic decomposition outweighs any potential benefits from isolating rapid responses. Among the considered image encodings, scalograms achieve the best results and the highest robustness, while recurrence plots provide the next-best accuracy but with a notably larger spread across runs. Markov transition fields, Gramian angular summation and difference fields remain clearly below scalograms in terms of performance in this setting. Across backbone families, ResNetV2 and EfficientNetV2 show the strongest overall results, whereas ConvNeXt underperforms across sizes. MobileNetV3 remains competitive despite substantially fewer parameters.

Overall, the strongest and most reliable performance is obtained when using scalograms of the full EDA signal together with pretrained CNN backbones from the ResNetV2 or EfficientNetV2 families under a locked fine-tuning configuration. Additionally, the scaling model capacity is family-dependent: ResNet101V2 improves modestly over ResNet50V2, EfficientNetV2-L does not improve over EfficientNetV2-M, and ConvNeXt scaling is not monotonic. These findings suggest that architecture choice and representation quality matter more than parameter count alone for EDA-based stress classification with time series image encodings.

Future work shall investigate further image encoding types focusing on different signal characteristics, ensemble models as well as other architectural groups, beyond CNNs. Furthermore, using more than one biophysiological signal as input in a multimodal approach will be explored in the future.

References

[1] D. Hellhammer, A. Stone, J. Hellhammer, and J. Broderick, "Measuring stress," *Encyclopedia of behavioral neuroscience*, vol. 2, pp. 186–191, 2010.

[2] R. A. Ross, S. L. Foster, and D. F. Ionescu, "The role of chronic stress in anxious depression," *Chronic Stress*, vol. 1, p. 2470547016689472, 2017, pMID: 32440578. [Online]. Available: https://doi.org/10.1177/2470547016689472

[3] N. Malykhin, J. Serrano, W. Pietrasik, and K. Hegadoren, "Effects of duration and intensity of psychological stressors on mental health outcomes," *Journal of Psychiatric Research*, vol. 187, pp. 211–222, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0022395625003024

[4] S. A. Kivimäki M., "Effects of stress on the development and progression of cardiovascular disease," *Nat Rev Cardiol.*, vol. 15, no. 4, pp. 215–229, Apr 2018.

[5] K. L. Tamashiro, R. R. Sakai, C. A. Shively, I. N. Karatsoreos, and L. P. Reagan, "Chronic stress, metabolism, and metabolic syndrome," *Stress*, vol. 14, no. 5, pp. 468–474, 2011, pMID: 21848434. [Online]. Available: https://doi.org/10.3109/10253890.2011.606341

[6] T. Iqbal, A. Elahi, P. Redon, P. Vazquez, W. Wijns, and A. Shahzad, "A review of biophysiological and biochemical indicators of stress for connected and preventive healthcare," *Diagnostics (Basel)*, vol. 11(3):556, 2021.

[7] J. Aguiló, P. Ferrer-Salvans, A. García-Rozo, A. Armario, A. Corbí, F. J. Cambra, R. Bailón, A. González-Marcos, G. Caja, S. Aguiló', R. López-Antón, A. Arza-Valdés, and J. M. Garzón-Rey, "Project ES3: attempting to quantify and measure the level of stress," *Revista de neurologia*, vol. 61, no. 9, p. 405—415, November 2015. [Online]. Available: https://doi.org/10.33588/rn.6109.2015136

[8] A.-A. Journault, A. Naud, and S. Lupien, "Gender and age differences in stress mindsets, a descriptive study from a large canadian sample," *Psychoneuroendocrinology*, vol. 160, p. 106688, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306453023006662

[9] C. H. Lindquist and C. A. Lindquist, "Gender differences in distress: mental health consequences of environmental stress among jail inmates," *Behav Sci Law.*, vol. 15, no. 4, pp. 503–23, 1997.

[10] H. D. Critchley, "Review: Electrodermal responses: What happens in the brain," *The Neuroscientist*, vol. 8, no. 2, pp. 132–142, 2002, pMID: 11954558. [Online]. Available: https://doi.org/10.1177/107385840200800209

[11] J. J. Braithwaite, D. P. Z. Watson, R. O. Jones, and M. A. Rowe, "Guide for analysing electrodermal activity & skin conductance responses for psychological experiments," *CTIT technical reports series*, 2013. [Online]. Available: https://api.semanticscholar.org/CorpusID:112717780

[12] X. Yu, J. Lu, W. Liu, Z. Cheng, and G. Xiao, "Exploring physiological stress response evoked by passive translational acceleration in healthy adults: a pilot study utilizing electrodermal activity and heart rate variability measurements," *Scientific Reports*, vol. 14, no. 1, p. 11349, May 2024. [Online]. Available: https://doi.org/10.1038/s41598-024-61656-5

[13] A. Greco, G. Valenza, J. Lázaro, J. M. Garzón-Rey, J. Aguiló, C. de la Cámara, R. Bailón, and E. P. Scilingo, "Acute stress state classification based on electrodermal activity modeling," *IEEE Transactions on Affective Computing*, vol. 14, no. 1, pp. 788–799, 2023.

[14] M. Kyrou, I. Kompatsiaris, and P. C. Petrantonakis, "Deep learning approaches for stress detection: A survey," *IEEE Transactions on Affective Computing*, vol. 16, no. 2, pp. 499–517, 2025.

[15] Z. Wang and T. Oates, "Encoding time series as images for visual inspection and classification using tiled convolutional neural networks," *In Workshops at the 29th AAAI Conference on Artificial Intelligence.*, 01 2015.

[16] P. Zontone, A. Affanni, A. Piras, and R. Rinaldo, "Convolutional neural networks using scalograms for stress recognition in drivers," in *2023 31st European Signal Processing Conference (EUSIPCO)*, 2023, pp. 1185–1189.

[17] S. Yang, Y. Gao, Y. Zhu, L. Zhang, Q. Xie, X. Lu, F. Wang, and Z. Zhang, "A deep learning approach to stress recognition through multimodal physiological signal image transformation," *Scientific Reports*, vol. 15, no. 1, p. 22258, 2025.

[18] Y. Hasanpoor, B. Tarvirdizadeh, K. Alipour, and M. Ghamari, "Multimodal signal fusion for stress detection using deep neural networks: a novel approach for converting 1d signals to unified 2d images," *Signal Image Video Process.*, vol. 19, no. 13, p. 1129, 2025. [Online]. Available: https://doi.org/10.1007/s11760-025-04734-z

[19] J. Lee, H. Lee, and M. Shin, "Driving stress detection using multimodal convolutional neural networks with nonlinear representation of short-term physiological signals," *Sensors*, vol. 21, no. 7, 2021. [Online]. Available: https://www.mdpi.com/1424-8220/21/7/2381

[20] M. Çitaku, L. Zott, S. Meyer-Nieberg, C. Küsel, and M. Hofmann, "Images of stress: Investigating image encodings and transfer learning for stress detection," in *2025 IEEE Symposium on Computational Intelligence in Health and Medicine (CIHM)*, 2025, pp. 1–7.

[21] P. Schmidt, A. Reiss, R. Duerichen, C. Marberger, and K. Van Laerhoven, "Introducing WESAD, a multimodal dataset for wearable stress and affect detection," in *Proceedings of the 20th ACM International Conference on Multimodal Interaction*, ser. ICMI '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 400–408. [Online]. Available: https://doi.org/10.1145/3242969.3242985

[22] M. Vetterli and J. Kovacevic, "Wavelets and subband coding," in *Prentice Hall Signal Processing Series*, 2013. [Online]. Available: https://api.semanticscholar.org/CorpusID:10506924

[23] J.-P. Eckmann, S. O. Kamphorst, and D. Ruelle, "Recurrence plots of dynamical systems," *Europhysics Letters*, vol. 4, no. 9, p. 973, nov 1987. [Online]. Available: https://doi.org/10.1209/0295-5075/4/9/004

[24] H. Poincaré, "Sur la probleme des trois corps et les équations de la dynamique," *Acta Mathematica*, vol. 13, no. 1-2, pp. 1–271, 1890. [Online]. Available: https://projecteuclid.org/journals/acta-mathematica/volume-13/issue-1-2

[25] F. Chollet *et al.*, "Keras Applications," https://keras.io/api/applications, 2026, accessed: 2026-01-30.

[26] A. Howard, M. Sandler, B. Chen, W. Wang, L.-C. Chen, M. Tan, G. Chu, V. Vasudevan, Y. Zhu, R. Pang, H. Adam, and Q. Le, "Searching for MobileNetV3," in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019, pp. 1314–1324.

[27] M. Tan and Q. Le, "EfficientNetV2: Smaller models and faster training," in *Proceedings of the 38th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, M. Meila and T. Zhang, Eds., vol. 139. PMLR, 18–24 Jul 2021, pp. 10 096–10 106. [Online]. Available: https://proceedings.mlr.press/v139/tan21a.html

[28] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *European conference on computer vision*. Springer, 2016, pp. 630–645.

[29] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, "A ConvNet for the 2020s," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022, pp. 11 966–11 976.

[30] V. Ashwin, R. Jegan, and P. Rajalakshmy, "Stress detection using wearable physiological sensors and machine learning algorithm," in *2022 6th International Conference on Electronics, Communication and Aerospace Technology*, 2022, pp. 972–977.

[31] J. J. Braithwaite, D. G. Watson, R. Jones, and M. Rowe, "A guide for analysing electrodermal activity (EDA) & skin conductance responses (SCRs) for psychological experiments," *Psychophysiology*, vol. 49, no. 1, pp. 1017–1034, 2013.

[32] A. Greco, G. Valenza, A. Lanata, E. P. Scilingo, and L. Citi, "cvxEDA: A convex optimization approach to electrodermal activity processing," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 4, pp. 797–804, 2015.

[33] J. Muradeli, "ssqueezepy," *GitHub. Note: https://github.com/OverLordGoldDragon/ssqueezepy*, 2020, accessed: 2026-01-30.

[34] J. Faouzi and H. Janati, "pyts: A python package for time series classification," *Journal of Machine Learning Research*, vol. 21, no. 46, pp. 1–6, 2020. [Online]. Available: http://jmlr.org/papers/v21/19-763.html

[35] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[36] O. N. Rahma, A. P. Putra, A. Rahmatillah, Y. S. K. A. Putri, N. D. Fajriaty, K. Ain, and R. Chai, "Electrodermal activity for measuring cognitive and emotional stress level," *Journal of Medical Signals & Sensors*, vol. 12, no. 2, pp. 155–162, 2022.

[37] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.