Four Stress Phenotypes Revealed Through Multimodal Thermal Imaging: The FenoStressNet Framework

Katarzyna Baran (k.baran@pollub.pl)[1]
[1]Department of Computer Science, Lublin University of Technology

Abstract

Stress is typically conceptualized as varying along a single intensity dimension, yet individuals with identical "high stress" labels often exhibit divergent responses. This work challenges this unidimensional model through FenoStressNet, a multimodal framework integrating facial thermal imaging, physiological monitoring, psychological assessment, and explainable AI. Analyzing 120 participants across controlled stress-induction tasks (19,200 thermal images), this study identifies four reproducible stress phenotypes that cross-cut traditional intensity categories: Cognitive-Dominant (elevated cognitive appraisal, moderate physiology), Physiological-Reactive (rapid autonomic arousal), Integrated-Responder (synchronized multimodal activation), and Discordant-Denier (high physiological response with low subjective awareness). These phenotypes exhibit distinct facial thermal patterns and differential weighting of cognitive, affective, and physiological components. These findings suggest a paradigm shift from intensity-based to phenotype-informed stress assessment, with significant implications for personalized interventions and cognitive theories of emotion.

Keywords: stress phenotypes, thermal imaging, explainable AI, multimodal fusion, cognitive-affective mechanisms

Introduction

Stress represents a complex interplay between cognitive appraisal, affective experience, and physiological arousal (Barrett & Simmons, 2015; Lazarus & Folkman, 1984). Despite its multidimensional nature, both clinical practice and computational modeling predominantly conceptualize stress as a unidimensional construct varying primarily in intensity (Epel et al., 2018). This approach does not capture heterogeneous responses: individuals labeled as experiencing "high stress" may instead exhibit cognitive rumination, autonomic hyperarousal, or dissociated states with minimal subjective awareness (Can et al., 2020). Recent advances in affective computing enable fine-grained, multimodal measurement (Garcia-Ceja et al., 2018). In particular, facial thermal imaging provides non-invasive insight into autonomic nervous system activity through blood flow changes in facial vasculature (Engert et al., 2014; Ioannou, Gallese, & Merla, 2014). However, existing studies typically: (1) employ small samples (N<50), limiting generalizability (Jaramillo-Quintanar et al., 2024); (2) use architectures pre-trained on natural RGB images (e.g., ImageNet), despite a significant domain mismatch with thermal physiology (Raghu et al., 2019); and (3) focus on classification accuracy rather than the discovery of latent response structures (Baran, 2025). Building on prior stress classification work using smartphone thermal cameras (Baran, 2024; 2025), this paper extends thermal stress recognition toward a cognitively grounded, phenotype-oriented framework.

This paper introduces FenoStressNet, a comprehensive framework designed to uncover stress heterogeneity. The contributions are: (1) discovery of four reproducible stress phenotypes through multimodal analysis of 120 participants; (2) development of a specialized neural architecture optimized for thermal facial analysis without ImageNet transfer; (3) integration of explainable AI (Grad-CAM, SHAP) to reveal phenotype-specific mechanisms; and (4) validation of phenotypes against behavioral and performance measures.

Methods

Participants and Protocol

The study included 120 healthy adult participants, consisting of 60 females and 60 males. The mean age was 28.4 years (SD = 5.7), with an age range from 22 to 45 years. Only individuals without a history of psychiatric or neurological disorders were eligible for inclusion. Additional exclusion criteria included cardiovascular diseases and facial skin conditions that could potentially affect thermal emission and reduce the accuracy of facial thermographic measurements.

Stress Induction Protocol

The stress induction procedure comprised four consecutive 15-minute stages designed to elicit progressively increasing levels of stress intensity. The protocol began with a baseline relaxation phase, during which no stressors were introduced and participants were instructed to rest quietly. This stage served as a physiological reference point and was characterized by a heart rate below 65 beats per minute and a subjective stress rating of 2 or lower on a self-report scale.

The second stage involved a low-intensity stress task based on serial subtraction of seven (Serial-7 arithmetic). This cognitively demanding activity was intended to induce mild stress and was associated with an increase in heart rate of approximately 10–15% relative to baseline, alongside self-reported stress levels ranging from 3 to 4. In the third stage, participants were asked to prepare a speech, a task commonly used to evoke anticipatory stress. This phase targeted a

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

medium level of stress intensity and was defined by a heart rate increase of 15–20% compared to baseline, as well as self-reported stress ratings between 5-6. The final stage consisted of competitive puzzle-solving tasks designed to elicit high levels of stress through time pressure and performance comparison. This stage was characterized by a heart rate increase exceeding 20% relative to baseline and subjective stress ratings of 7 or higher, indicating a high-intensity stress response.

Ground truth labels were established for each 5-minute epoch using a multimodal consensus approach integrating subjective, physiological, and expert-based assessments. Participants provided self-reported stress ratings on a standardized 0–10 scale, capturing perceived stress levels during each epoch. Objective physiological thresholds, primarily changes in heart rate relative to baseline, were applied to quantify autonomic stress responses. Additionally, all epochs were independently evaluated by two trained expert coders, whose assessments demonstrated high inter-rater reliability (Cohen's $\kappa = 0.82$), indicating strong agreement. The final ground truth label for each epoch was determined through consensus across these three sources.

Given that each of the 120 participants completed four experimental stages, the procedure yielded a total of 480 labeled epochs used for subsequent analysis.

Multimodal Data Acquisition

Thermal data were acquired using an Infiray TS2+ thermal camera (resolution 256 × 192 pixels, spectral range 8–14 µm) mounted on an iPhone 14 Pro. Thermal recordings were captured at a sampling rate of 1 Hz across all experimental stages. This acquisition protocol yielded a total of 19,200 thermal images, corresponding to 120 participants completing four experimental stages, with approximately 40 thermal frames collected per stage.

Physiological measurements were collected concurrently with thermal imaging to provide complementary objective indicators of stress responses. Heart rate was continuously recorded at 1 Hz using a Polar H10 chest strap. Blood pressure was measured using an Omron M7 automated sphygmomanometer immediately before and after each experimental stage to assess stage-specific cardiovascular changes. Additionally, electrodermal activity was optionally recorded in a subset of participants using an Empatica E4 wearable device at a sampling rate of 4 Hz, enabling assessment of sympathetic nervous system activity.

Following completion of the experimental tasks, participants completed a set of standardized psychological questionnaires to assess perceived stress levels and coping strategies. These included the 10-item Perceived Stress Scale (PSS-10) (Cohen et al., 1983; Cohen & Williamson, 1988), the Coping Inventory for Stressful Situations (CISS) (Endler & Parker, 1990), and the brief MINICOPE coping questionnaire (Brambila-Tapia et al., 2023).

Thermal and physiological data were processed using a standardized preprocessing pipeline prior to analysis. Facial regions were automatically detected using the MediaPipe Face Mesh framework, enabling robust identification of facial landmarks across thermal frames. Regions of interest were extracted based on 68-landmark facial alignment and cropped to a standardized resolution of 224 × 224 pixels. Thermal values were normalized on a per-participant basis using min–max scaling within a temperature range of 28–36°C to account for inter-individual variability.

To minimize motor and speech-related artifacts in thermal analysis, two controls were implemented. First, frames with detected facial landmark velocity exceeding 5 pixels per frame (computed via MediaPipe landmark trajectories) were excluded from analysis (3.8% of all frames). Second, to exclude perioral thermal artifacts associated with speech production, Grad-CAM activation maps were recomputed after masking the mouth and chin regions (lower 30% of the face). Phenotype-specific thermal patterns (e.g., forehead focus for Cognitive-Dominant, nasal/cheek focus for Physiological-Reactive) remained highly correlated with full-face patterns (Pearson's $r > 0.87$), confirming that these patterns reflect autonomic stress responses rather than motor or speech artifacts. Physiological signals were temporally aligned with thermal frames using a 5-second rolling average to ensure synchronization across modalities.

Table 1 FenoStressNet architecture specifications

Layer	Type	Output size	Kernel/Stride	Params	Special Components
Input	Thermal ROI	1×224×224	-	-	Single-channel thermal image
Stem	Conv2D	16 ×112×112	3×3 / 2	160	ReLU, BatchNorm
Stage1	GhostModule	32×112×112	3×3 / 1	1 024	Coordinate Attention
	DynamicPath	32×112×112	-	640	Weight learning
Stage2	GhostModule	64×56×56	3×3 / 2	4 096	Coordinate Attention
	DynamicPath	64×56×56	-	2 560	Weight learning
Stage3	GhostModule	128×28×28	3×3 / 2	16 384	Coordinate Attention
	DynamicPath	128×28×28	-	10 240	Weight learning
GAP	Global Avg Pool	128	-	-	-
Fusion	Concatenate	134	-	-	Physiological+psych
Classifier	FC → Softmax	4	-	540	Output: none/low/med/high
Total	**-**		**-**	**35 444**	**-**

FenoStressNet Architecture

The proposed FenoStressNet architecture is a lightweight convolutional neural network designed specifically for thermal facial stress recognition. A detailed layer-by-layer specification of the model is provided in Table 1. The network takes as input a single-channel thermal facial region of interest (ROI) with a spatial resolution of 224 × 224 pixels. An initial convolutional stem consisting of a 3 × 3 Conv2D layer with stride 2 performs early spatial downsampling and low-level feature extraction, followed by batch normalization and ReLU activation. The core of the network comprises three sequential stages built using Ghost Modules, which progressively increase the number of feature channels from 32 to 128 while reducing spatial resolution. Each Ghost Module employs lightweight linear operations to generate additional feature maps at a significantly reduced computational cost compared to standard convolutions. To enhance spatial representational capacity, Coordinate Attention mechanisms are integrated within each stage, enabling the network to capture long-range dependencies in thermal gradients along both horizontal and vertical axes. Between the main convolutional stages, Dynamic Path Weighting blocks are introduced to adaptively learn the relative importance of multi-scale features. These blocks employ a gating mechanism that assigns trainable weights to different feature paths, allowing the model to dynamically emphasize the most informative representations during training. Following the final convolutional stage, global average pooling is applied to aggregate spatial features into a compact 128-dimensional embedding. This representation is then fused with six normalized physiological and psychological features via late multimodal fusion implemented through feature concatenation. The resulting joint feature vector is passed to a fully connected classifier with a softmax output layer, producing four discrete stress-level predictions: none, low, medium, and high.

The efficiency of FenoStressNet is largely driven by the use of Ghost Modules, which replace conventional convolutional layers with computationally inexpensive linear operations, achieving an approximate 40% reduction in parameter count while preserving feature diversity (Han et al., 2020). Coordinate Attention modules further enhance discriminative performance by embedding positional information into channel-wise attention, which is particularly effective for capturing spatially distributed thermal patterns associated with stress responses (Hou et al., 2021). Dynamic Path Weighting enables adaptive feature aggregation across multiple scales by learning optimal combination weights through a gating mechanism (Li et al., 2023). Finally, multimodal fusion is performed at a late stage, where deep thermal features are concatenated with complementary physiological and psychological signals to improve robustness and generalization.

The network was trained using the AdamW optimizer with an initial learning rate of 0.001, momentum parameters $\beta_1 = 0.9$ and $\beta_2 = 0.999$, and a weight decay of 0.01. A cosine annealing learning rate schedule with warm restarts was used to ensure stable convergence. Training was performed with a batch size of 32, balanced across stress classes, for a maximum of 200 epochs, with early stopping applied if validation performance did not improve for 20 consecutive epochs. To address class imbalance, Focal Loss with a focusing parameter $\gamma = 2.0$ was used. All experiments were conducted on a single NVIDIA RTX 4090 GPU.

Phenotype Discovery Pipeline

Importantly, stress intensity classification was used solely as a supervisory signal for learning a robust multimodal representation, while phenotype discovery was performed independently through unsupervised analysis of the learned feature space. For each thermal image, deep feature representations were extracted from the penultimate layer of the trained FenoStressNet model. This process produced fixed-length 134-dimensional feature vectors capturing high-level abstractions of facial thermal patterns associated with stress-related physiological responses. These embeddings formed the basis for subsequent explainability, visualization, and clustering analyses.

To interpret the contribution of individual feature modalities to the model's predictions, feature importance was assessed using the KernelSHAP framework. SHAP values were computed separately for thermal, physiological, and psychological feature subsets, enabling modality-specific attribution of predictive influence. This approach provided a model-agnostic explanation of how different input features contributed to stress-level classification decisions.

High-dimensional feature embeddings were projected into a two-dimensional space using t-distributed stochastic neighbor embedding (t-SNE) to facilitate visualization and exploratory analysis. Importantly, clustering was performed on the original high-dimensional embeddings, while t-SNE was used solely for visualization. The t-SNE algorithm was configured with a perplexity value of 30 and run for 1,000 iterations, balancing local and global structure preservation in the reduced feature space.

Unsupervised clustering was performed on the original high-dimensional feature embeddings using the HDBSCAN algorithm. Clustering parameters were set to a minimum cluster size of 15 and a minimum of 5 samples per core point, enabling the identification of dense, well-separated clusters while automatically labeling noise and outlier samples. This density-based approach allowed the discovery of latent stress-related phenotypes without imposing a predefined number of clusters.

To validate the optimal number of clusters, the analysis systematically evaluated k=2 through k=6 using silhouette scores and bootstrap stability analysis on the learned 134-dimensional embeddings. k=4 achieved the highest silhouette score (0.67), compared to k=3 (0.58), k=5 (0.61), and k=6 (0.59). The Davies–Bouldin index was also lowest for k=4 (0.54) versus k=3 (0.68), k=5 (0.61), and k=6 (0.63). Bootstrap stability analysis (1,000 iterations, 80% subsampling with replacement) confirmed that 4-cluster solutions emerged in 92% of runs, while alternative k values showed lower cluster stability (<75% agreement across bootstrap samples). Alternative clustering algorithms (Gaussian Mixture Models, DBSCAN with varied epsilon parameters) converged on four clusters, supporting the robustness of this decomposition. The resulting clusters were externally validated by comparing them across independent psychological and physiological outcome measures. One-way analysis of variance (ANOVA) was conducted to assess statistically significant differences between clusters, providing evidence for the construct validity and interpretability of the identified stress-related groupings.

Model performance for stress-level classification was evaluated using multiple standard metrics to capture both overall and class-specific predictive quality. Accuracy was calculated as the proportion of correctly classified samples across all classes. To account for potential class imbalance, precision, recall, and F1-score were computed using macro-averaging, ensuring that each stress level contributed equally to the overall metric regardless of its prevalence in the dataset.

The quality of unsupervised clustering results was assessed using density- and distance-based metrics. The silhouette score quantified how similar each sample was to its own cluster compared to other clusters, providing an indication of cluster cohesion and separation. The Davies–Bouldin index measured the average similarity between clusters, with lower values indicating better-defined, more compact clusters.

To evaluate relationships between categorical variables, chi-square (χ^2) tests for independence were performed. For continuous variables, both Pearson and Spearman correlation coefficients were calculated depending on the distribution and linearity of the data. These statistical analyses supported the validation of model outputs and cluster assignments against independent psychological and physiological measures, providing a comprehensive assessment of construct validity.

Results

Stress Intensity Classification (Representation Learning Stage)

Table 2 and Figure 1 summarize the performance of FenoStressNet compared to several baseline models on the test set, which included 12 participants and a total of 1,920 thermal images. FenoStressNet achieved the highest overall accuracy (87.3%) and macro-averaged F1-score (86.9%) while maintaining a compact model size (35,440 parameters) and fast inference time (14.2 ms per image). In comparison, conventional deep learning models such as MobileNetV2, ResNet-18, and EfficientNet-B0 exhibited lower classification performance and substantially larger parameter counts, while an SVM trained on handcrafted features showed the lowest accuracy (71.2%).

The accompanying confusion matrix illustrates that FenoStressNet consistently distinguished between the four stress levels, achieving recall values above 86% for all classes and balanced precision across categories. These results demonstrate that the proposed architecture provides both high accuracy and computational efficiency, making it suitable for real-time stress intensity monitoring using thermal imaging.

An ablation study was also performed, the results are shown in Table 3.

	Predicted →				
Actual ↓	**None**	**Low**	**Medium**	**High**	**Recall**
None	141	7	3	1	92.8%
Low	6	132	9	3	88.0%
Medium	4	8	129	9	86.0%
High	2	3	10	135	90.0%
Precision	92.2%	88.0%	85.4%	91.2%	

Figure 1: Confusion Matrix for FenoStressNet

Discovered Stress Phenotypes

In this work, stress phenotypes are defined as recurrent, intensity-independent patterns of multimodal response organization, rather than as levels of stress magnitude. Table 4 summarizes the four stress phenotypes identified through unsupervised clustering of thermal and multimodal features.

Each phenotype is characterized by dominant modalities (based on SHAP values), distinctive thermal activation patterns, and psychological profiles. The Cognitive-Dominant group (34.2% of participants) is primarily influenced by self-reported measures, particularly PSS-10 and CISS, with thermal activity concentrated on the forehead and psychological tendencies toward high rumination and low avoidance. The Physiological-Reactive phenotype (28.3%) is dominated by heart rate and blood pressure, with elevated thermal responses in the nasal and cheek regions, and participants showing low self-awareness but strong task-focused engagement. Integrated-Responder participants (21.7%) exhibit a balanced contribution of all modalities, diffuse whole-face thermal patterns, and adaptive coping with high awareness. Finally, the Discordant-Denier group (15.8%) is primarily driven by physiological signals, shows thermal activity around the chin and neck, and reports low perceived stress despite measurable physiological reactivity, reflecting a denial-based coping style.

Table 2: Macro-averaged classification performance (test set: 12 participants, 1,920 images)

Model	Accuracy	Precision	Recall	F1-Score	Params
FenoStressNet	**87.3%**	**86.8%**	**87.1%**	**86.9%**	**35 444**
MobileNetV2	82.1%	81.7%	81.9%	81.8%	3 504 872
ResNet-18	79.4%	78.9%	79.1%	79.0%	11 689 512
EfficientNet-B0	84.2%	83.8%	84.0%	83.9%	5 288 548
SVM (handcrafted)	71.2%	70.5%	70.8%	70.6%	-

Table 3: Component ablation (5-fold cross-validation)

Configuration	Accuracy	Δ vs Full	Parameters	Interpretation
Full FenoStressNet	87.3±1.2%	-	35 444	Baseline
w/o Ghost Modules	83.1±1.5%	-4.2%	58 912	Ghost reduces params 40%
w/o Coord Attention	84.6±1.3%	-2.7%	33 120	Spatial attention critical
w/o Dynamic Paths	85.2±1.4%	-2.1%	34 880	Multi-scale fusion helps
ImageNet Pretrained (replicated to 3-channel)	79.8±2.1%	-7.5%	35 444	Thermal–RGB domain mismatch

These phenotypes highlight the heterogeneity of stress responses: participants with similar overall stress intensity can display markedly different combinations of physiological, thermal, and psychological profiles, underscoring the value of multimodal assessment for capturing nuanced stress-related patterns.

Table 5 presents the distribution of the four stress phenotypes across the four stress intensity levels. The cross-tabulation shows that each phenotype appears across multiple intensity levels, highlighting the heterogeneity of stress responses. For example, the Physiological-Reactive and Discordant-Denier phenotypes are more prevalent at higher stress levels, whereas Cognitive-Dominant and Integrated-Responder participants are distributed more evenly across intensity conditions. This pattern supports the notion that stress intensity alone does not fully capture individual differences in physiological, thermal, and psychological responses.

Table 4: Four Stress Phenotypes with Characteristics

Phenotype	N (%)	Dominant Modality (mean normalized SHAP contribution)	Thermal Pattern (Grad-CAM)	Psychological Profile
Cognitive-Dominant	41 (34.2%)	PSS-10 (0.42), CISS (0.38)	Forehead focus (ΔT=+1.2±0.3°C)	High rumination, low avoidance
Physiological-Reactive	34 (28.3%)	HR (0.45), BP (0.35)	Nasal/cheek (ΔT=+1.8±0.4°C)	Low awareness, task-focused
Integrated-Responder	26 (21.7%)	All balanced (~0.25 each)	Whole-face diffuse	Adaptive coping, high awareness
Discordant-Denier	19 (15.8%)	HR (0.50), BP (0.30)	Chin/neck (ΔT=+1.5±0.5°C)	Denial coping, low self-report

Table 5: Phenotype x Intensity cross-tabulation

Phenotype ↓ / Intensity →	None	Low	Medium	High	Total
Cognitive-Dominant	15 (36.6%)	12 (29.3%)	9 (22.0%)	5 (12.2%)	41
Physiological-Reactive	5 (14.7%)	8 (23.5%)	10 (29.4%)	11 (32.4%)	34
Integrated-Responder	8 (30.8%)	6 (23.1%)	7 (26.9%)	5 (19.2%)	26
Discordant-Denier	2 (10.5%)	4 (21.1%)	4 (21.1%)	9 (47.4%)	19
Total	**30**	**30**	**30**	**30**	**120**

Phenotype assignments showed moderate temporal stability across the four sequential stress stages (baseline, low, medium, high). Across all participants, 78 of 120 (65.0%) maintained the same phenotype classification throughout all four stages (Cohen's $\kappa = 0.62$, 95% CI [0.55-0.69]), suggesting that phenotypes capture trait-like response patterns rather than task-specific states. Discordant-Denier showed the highest stability (14 of 19, 73.7%), while Integrated-Responder showed the lowest (15 of 26, 57.7%).

Validation Against External Measures

Phenotype validation using external measures confirmed meaningful differences in stress-related behavior and psychological profiles. Cognitive-Dominant participants exhibited relatively high task errors (12.3 ± 3.2) and elevated rumination scores (4.1 ± 0.5), whereas Physiological-Reactive participants showed the highest task errors (18.7 ± 4.1) but lower rumination (2.8 ± 0.6). Integrated-Responder individuals displayed intermediate task errors (15.2 ± 3.5) with the highest interoceptive accuracy (0.75 ± 0.07) and moderate rumination (3.2 ± 0.5). Discordant-Denier participants had task errors of 17.9 ± 3.8, the lowest interoceptive accuracy (0.48 ± 0.11), and rumination scores of 2.9 ± 0.6. Recovery times also varied across phenotypes, with Cognitive-Dominant at 8.4 ± 2.1 min, Physiological-Reactive at 5.2 ± 1.8 min, Integrated-Responder at 6.8 ± 2.0 min, and Discordant-Denier at 7.9 ± 2.3 min. ANOVA tests confirmed significant differences across phenotypes for all measures ($F(3,116)$ = 6.2–22.1, $p < 0.01$), supporting the construct validity of the clusters.

Discussion

Crucially, the identified stress phenotypes are not derived from stress intensity labels but emerge from unsupervised structure in the learned multimodal representation space. FenoStressNet demonstrates that specialized architectures outperform repurposed ImageNet models for thermal physiology. The 7.5% accuracy gap between from-scratch training and ImageNet transfer aligns with findings in medical imaging (Raghu et al., 2019; Albertetti et al., 2020), highlighting the importance of domain-matched design. The Ghost Module implementation achieved 40% parameter reduction without performance loss—critical for eventual clinical/field deployment.

The identification of four reproducible stress phenotypes challenges purely intensity-based models of stress and highlights systematic individual differences in stress responding. From a cognitive science perspective, these findings are consistent with constructivist and predictive models of emotion, in which stress responses reflect organized patterns of appraisal, bodily signaling, and awareness rather than a single arousal continuum. The four phenotypes map onto specific theoretical mechanisms. Cognitive-Dominant aligns with the perseverative cognition hypothesis (Brosschot et al., 2006), where sustained cognitive activity prolongs physiological activation. Physiological-Reactive corresponds to autonomic specificity models (Kreibig, 2010), characterized by rapid sympathetic engagement without cognitive elaboration. Discordant-Denier exemplifies interoceptive prediction errors within predictive coding frameworks (Barrett & Simmons, 2015), where bodily states are misrepresented at the metacognitive level. Integrated-Responder reflects optimal allostatic regulation (McEwen, 2017), with balanced coordination across systems. First, the *Discordant-Denier* phenotype may represent a potentially relevant risk pattern, characterized by elevated physiological arousal despite low subjective awareness, suggesting that self-report measures alone may fail to detect vulnerable individuals (Critchley et al., 2005). Second, *Cognitive-Dominant* stress is associated with pronounced prefrontal thermal activation and prolonged recovery, which may indicate that cognitive-based interventions could be more effective than relaxation-focused approaches for this group (Barrett & Simmons, 2015; Epel et al., 2018). Finally, all phenotypes span multiple stress intensity levels, demonstrating that stress cannot be fully characterized by magnitude alone but requires consideration of the qualitative organization of physiological, thermal, and psychological responses.

The observed differences in thermal patterns and modality weightings across phenotypes suggest distinct underlying physiological and cognitive processes. Cognitive-Dominant individuals exhibit prefrontal thermal activation, reflecting engagement in appraisal and rumination processes, with slower recovery times indicating sustained cognitive involvement. Physiological-Reactive participants display rapid nasal and cheek warming, consistent with parasympathetic withdrawal and sympathetic activation, with minimal cortical engagement. The Discordant-Denier phenotype shows thermal activity focused on the chin and neck coupled with low interoceptive accuracy, pointing to potential interoceptive dysfunction or dissociation. Finally, Integrated-Responder individuals demonstrate balanced multimodal responses, suggesting adaptive and flexible stress processing across physiological, thermal, and psychological domains. Our healthy sample limits generalizability to clinical populations. Future work should validate phenotypes in anxiety, depression, and PTSD cohorts. Longitudinal tracking could reveal phenotype stability, transitions, and treatment response differences. Neuroimaging correlation (fMRI with thermal imaging) could identify neural substrates of each phenotype.

Conclusion

This paper presented FenoStressNet, a multimodal framework that discovers four distinct and reproducible stress phenotypes: Cognitive-Dominant, Physiological-Reactive, Integrated-Responder, and Discordant-Denier. These phenotypes cross-cut traditional intensity categories and exhibit distinct thermal patterns, modality weightings, and behavioral correlates. This work shifts stress assessment from a unidimensional intensity model to a phenotype-based approach, with significant implications for personalized interventions and cognitive theories of emotion. The FenoStressNet framework demonstrates how specialized computational architectures combined with explainable AI can reveal latent structures in complex psychological phenomena.

Acknowledgments

The research was conducted as part of the project "Human Stress Recognition Using Thermographic Imaging and Neural Networks" which received positive approval from the Ethics Committee for Scientific Research (approval No. 4/2024).

References

Albertetti, F., Simalastar, A., & Rizzotti-Kaddouri, A. (2020, December). Stress detection with deep learning approaches using physiological signals. In *International Conference on IoT Technologies for HealthCare* (pp. 95-111). Cham: Springer International Publishing.

Baran, K. (2025). StressIRNet: A Novel Lightweight CNN Architecture for Stress Classification Leveraging Smartphone Thermal Imaging Modality. *IEEE Access*, *13*, 194475-194488.

Baran, K. (2024). Fuzzy Ensemble Learning and Lightweight CNN for Stress Classification. In *ECAI 2024* (pp. 4741-4747). IOS Press.

Barrett, L. F., & Simmons, W. K. (2015). Interoceptive predictions in the brain. *Nature Reviews Neuroscience, 16*(7), 419-429. https://doi.org/10.1038/nrn3950

Brambila-Tapia, A. J. L., Martínez-Arriaga, R. J., González-Cantero, J. O., González-Becerra, V. H., Reyes-Domínguez, Y. A., Ramírez-García, M. L., & Macías-Espinoza, F. (2023, April). Brief COPE short version (Mini-COPE): A proposal of item and factorial reduction in mexican population. In *Healthcare* (Vol. 11, No. 8, p. 1070). MDPI.

Brosschot, J. F., Gerin, W., & Thayer, J. F. (2006). The perseverative cognition hypothesis: A review of worry, prolonged stress-related physiological activation, and health. *Journal of Psychosomatic Research*, 60(2), 113-124.

Can, Y. S., Chalabianloo, N., Ekiz, D., & Ersoy, C. (2019). Continuous stress detection using wearable sensors in real life: Algorithmic programming contest case study. *Sensors, 19*(8), 1849. https://doi.org/10.3390/s19081849

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior, 24*(4), 385-396.

Cohen, S., & Williamson, G. (1988). Perceived stress in a probability sample of the United States. In S. Spacapan & S. Oskamp (Eds.), *The Social Psychology of Health* (pp. 31-67). Sage.

Critchley, H. D., Elliott, R., Mathias, C. J., & Dolan, R. J. (2000). Neural activity relating to generation and representation of galvanic skin conductance responses: a functional magnetic resonance imaging study. *Journal of Neuroscience*, *20*(8), 3033-3040.

Endler, N. S., & Parker, J. D. (1990). Multidimensional assessment of coping: a critical evaluation. *Journal of personality and social psychology*, *58*(5), 844.

Engert, V., Merla, A., Grant, J. A., Cardone, D., Tusche, A., & Singer, T. (2014). Exploring the use of thermal infrared imaging in human stress research. *PloS one*, *9*(3), e90782.

Epel, E. S., Crosswell, A. D., Mayer, S. E., Prather, A. A., Slavich, G. M., Puterman, E., & Mendes, W. B. (2018). More than a feeling: A unified view of stress measurement for population science. *Frontiers in neuroendocrinology*, *49*, 146-169.

Garcia-Ceja, E., Riegler, M., Nordgreen, T., Jakobsen, P., Oedegaard, K. J., & Tørresen, J. (2018). Mental health monitoring with multimodal sensing and machine learning: A survey. *Pervasive and Mobile Computing*, *51*, 1-26.

Han, K., Wang, Y., Tian, Q., Guo, J., Xu, C., & Xu, C. (2020). Ghostnet: More features from cheap operations. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 1580-1589).

Hou, Q., Zhou, D., & Feng, J. (2021). Coordinate attention for efficient mobile network design. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 13713-13722).

Ioannou, S., Gallese, V., & Merla, A. (2014). Thermal infrared imaging in psychophysiology: potentialities and limits. *Psychophysiology*, *51*(10), 951–963. https://doi.org/10.1111/psyp.12243

Jaramillo-Quintanar, D., Gomez-Reyes, J. K., Morales-Hernandez, L. A., Dominguez-Trejo, B., Rodriguez-Medina, D. A., & Cruz-Albarran, I. A. (2023). Automatic segmentation of facial regions of interest and stress detection using machine learning. *Sensors*, *24*(1), 152.

Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: A review. *Biological Psychology*, 84(3), 394-421.

Lazarus, R. S. (1984). *Stress, appraisal, and coping* (Vol. 445). Springer.

Li, S., Mao, Y., Zhang, F., Wang, D., & Zhong, G. (2023). DLW-NAS: differentiable light-weight neural architecture search. *Cognitive Computation*, *15*(2), 429-439.

McEwen, B. S. (2017). Neurobiological and Systemic Effects of Chronic Stress. *Chronic Stress*. SAGE Publications Inc.

Raghu, M., Zhang, C., Kleinberg, J., & Bengio, S. (2019). Transfusion: Understanding transfer learning for medical imaging. *Advances in Neural Information Processing Systems, 32*, 3347-3357.