**JoulesEye: Energy Expenditure Estimation and Respiration Sensing from Thermal Imagery While Exercising**

**Authors :-** RISHIRAJ ADHIKARY, MAITE SADEH , NIPUN BATRA, MAYANK GOEL

**Summary :-**

**JoulesEye: Energy Expenditure Estimation and Respiration Sensing from Thermal Imagery While Exercising - Comprehensive 5000-Word Analysis**

**Introduction and Problem Statement**

The proliferation of smartphone and smartwatch technologies has fundamentally transformed personal health monitoring, creating unprecedented opportunities for real-time physiological data collection. These ubiquitous devices, equipped with increasingly sophisticated sensor arrays, provide users with immediate access to vital health metrics including heart rate, step count, activity levels, and various fitness parameters. However, beneath this technological advancement lies a critical limitation that undermines both clinical utility and practical application: the persistent and significant inaccuracy in energy expenditure estimation, commonly known as calorie burn measurement.

This fundamental flaw represents more than a technical inconvenience—it constitutes a substantial barrier to effective health management and evidence-based medical decision-making. The global obesity epidemic, affecting over one billion people worldwide, demands accurate metabolic monitoring tools for successful intervention and management. Current commercial devices demonstrate Mean Absolute Percentage Errors (MAPE) exceeding 37% when estimating energy expenditure compared to gold-standard measurements, rendering them unsuitable for applications requiring precision, clinical diagnosis, or medical intervention.

**JoulesEye**, a groundbreaking research initiative developed through collaboration between IIT Gandhinagar, Cornell University, and Carnegie Mellon University, represents a paradigmatic shift in wearable energy expenditure monitoring. Published in December 2023 in the Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, this innovative research introduces a thermal camera-based system that achieves unprecedented 5.8% MAPE compared to indirect calorimetry gold standards—representing more than six-fold improvement over existing commercial approaches.

The core innovation centers on thermal imaging technology to monitor respiration rate rather than relying exclusively on traditional heart rate measurements. This revolutionary approach challenges the conventional wisdom that has dominated wearable health technology development for decades. The physiological foundation underlying this advancement stems from respiration rate's superior correlation with metabolic activity, body composition variations, and real-time energy demands compared to heart rate alone.

The research addresses the fundamental problem that current consumer wearables suffer from oversimplified approaches to energy expenditure estimation. These devices, designed primarily for consumer appeal rather than clinical accuracy, depend heavily on heart rate data as the primary physiological indicator. This approach fails catastrophically when confronted with the complex reality of human metabolism, which varies dramatically across individuals based on body composition, fitness level, metabolic efficiency, genetic factors, and numerous physiological characteristics that heart rate measurements cannot capture.

Investigation revealed that exclusive reliance on heart rate creates systematic errors that become more pronounced across different user populations. Particularly concerning is the finding that Apple Watch errors can reach 51.8% in overweight individuals compared to 29.7% in normal-weight users, while JoulesEye maintains consistent 5.2-6.9% accuracy across all body compositions. This differential performance highlights fundamental limitations in heart rate-only approaches and demonstrates respiration rate's superior ability to account for individual physiological variations.

The motivation extends beyond immediate accuracy concerns to address the significant gap between clinical-grade measurement capabilities and consumer accessibility. Current gold-standard methods—including Double Labeled Water tests, direct calorimetry, and indirect calorimetry—require extensive clinical infrastructure, expensive equipment, trained personnel, and often necessitate participant confinement for extended periods. JoulesEye's thermal imaging approach potentially bridges this gap by bringing clinical-grade accuracy to consumer-accessible form factors.

The practical implications span critical applications in clinical medicine, sports performance optimization, and public health management. For athletes engaged in intensive training where precise energy expenditure measurement is crucial for performance optimization and injury prevention, current smartwatch errors exceeding 30% can lead to inadequate caloric intake, compromising training effectiveness and potentially causing health complications. The International Society of Sports Nutrition emphasizes that elite athletes may burn over 1200 calories per hour during intensive exercise, requiring corresponding caloric intake for optimal performance and health maintenance.

**Background and Related Work**

Understanding energy expenditure measurement requires comprehensive examination of the physiological processes underlying human metabolism and the technological approaches developed to quantify these processes. Total Energy Expenditure (TEE) represents the sum of all energy-consuming processes within the human body, decomposed into three distinct components with unique characteristics and measurement challenges.

**Resting Energy Expenditure (REE)** constitutes baseline energy consumption for fundamental physiological processes including cellular maintenance, protein synthesis, ion transport, and basic organ function. This component typically accounts for 60-75% of total energy expenditure in sedentary individuals and varies significantly based on age, gender, body size, body composition, ethnicity, physical fitness level, hormonal status, and genetic factors. The complexity becomes apparent when considering that individuals with identical body weights can have dramatically different REE values based on muscle-to-fat ratios, as muscle tissue is metabolically more active than adipose tissue.

**Diet-Induced Energy Expenditure (DEE)** represents the energy cost associated with food digestion, absorption, transport, metabolism, and nutrient storage. This component typically accounts for 8-10% of total energy expenditure and varies significantly based on macronutrient composition, meal timing, and individual metabolic efficiency. DEE measurement requires sophisticated equipment and controlled conditions, as effects must be measured over several hours following food consumption.

**Activity-Induced Energy Expenditure (EE)** represents the most variable component and constitutes the primary focus of fitness monitoring applications. This includes both planned exercise activities and non-exercise activity thermogenesis (NEAT), encompassing energy expenditure from daily activities including posture maintenance, spontaneous muscle contraction, and fidgeting. EE variability across individuals and activities makes it the most challenging component to measure accurately.

**Traditional Measurement Approaches** reveal fundamental trade-offs between accuracy and practicality that have shaped energy expenditure measurement technology development. The Double Labeled Water (DLW) test employs stable isotopes of hydrogen and oxygen to track metabolic water production over extended periods. While highly accurate, DLW requires specialized laboratory facilities, expensive isotopes, complex analytical procedures, and multiple clinical visits over 14-day periods, making it completely impractical for routine monitoring.

Direct calorimetry represents the most fundamental approach, directly quantifying heat production through precisely controlled chamber environments. Participants are confined within insulated chambers where every calorie of heat production is measured through sophisticated temperature monitoring systems. While theoretically perfect in accuracy, direct calorimetry requires specialized facilities costing millions of dollars and 24-48 hour subject confinement, making it unsuitable for realistic exercise scenarios.

Indirect calorimetry offers more practical alternatives by measuring oxygen consumption and carbon dioxide production to estimate energy expenditure through established metabolic equations. This method leverages the principle that energy metabolism involves fuel oxidation consuming oxygen and producing carbon dioxide in predictable ratios. The Fitmate Pro indirect calorimeter used in JoulesEye research employs this principle, converting VO2 measurements to energy expenditure using the established 5 kcal per liter conversion factor.

**Consumer Device Limitations** become apparent when examining physiological bases of current energy expenditure estimation approaches in commercial wearables. Heart rate-based estimation relies on relationships between cardiac output and oxygen delivery to working muscles, expressed as: EE = HR × SV × δavO2 × γ, where HR represents heart rate, SV denotes stroke volume, δavO2 represents arteriovenous oxygen difference, and γ is the oxygen-to-energy conversion coefficient.

This equation reveals fundamental weaknesses in heart rate-only approaches: both stroke volume and arteriovenous oxygen difference are highly variable parameters depending on body composition, fitness level, exercise intensity, hydration status, environmental conditions, and genetic factors affecting cardiac and vascular function. Stroke volume varies dramatically across individuals based on cardiac fitness, with trained athletes potentially having stroke volumes 50-100% higher than sedentary individuals at identical heart rates.

**The Respiration Rate Advantage** emerges from deeper understanding of physiological connections between breathing patterns, metabolic demands, and individual body composition characteristics. Medical literature demonstrates strong correlations between respiratory function and body composition, particularly adiposity levels, which directly affect metabolic efficiency and energy expenditure patterns.

The physiological basis lies in fundamental relationships between ventilation and metabolic demands. Unlike heart rate, which primarily reflects cardiovascular responses to oxygen demands, respiration rate reflects integrated responses of cardiovascular, respiratory, metabolic, and neurological systems. The brain's motor cortex serves as primary regulator of both respiration rate and aerobic capacity, creating direct neurological links between breathing patterns and energy expenditure absent in heart rate measurements.

**Thermal Imaging for Respiration Detection** represents significant technological advancement over previous respiration monitoring methods, addressing fundamental limitations preventing accurate respiration measurement during vigorous exercise. Traditional contact-based methods using chest belts suffer from severe motion artifacts during vigorous exercise, where oscillatory movements from running or cycling occur at frequencies similar to respiration rates, making it impossible to distinguish between movement artifacts and actual respiratory signals.

RGB camera-based methods fail during exercise due to external lighting variations and motion-induced artifacts that overwhelm subtle color changes associated with breathing. These methods rely on detecting minute skin color changes caused by blood volume variations during respiratory cycles, but during exercise, much larger color changes occur due to increased blood flow, sweating, and varying light conditions.

**Methodology and System Design**

The JoulesEye system architecture represents sophisticated integration of hardware components, signal processing algorithms, and machine learning models specifically designed to extract energy expenditure information from thermal imagery during vigorous exercise. The comprehensive approach addresses multiple technical challenges inherent in physiological monitoring during motion while maintaining accuracy standards necessary for practical deployment.

**Hardware Configuration and Integration** centers on the FLIR One Pro thermal imaging camera, a smartphone attachment providing 160×120 pixel resolution thermal video capture at 8.6 frames per second with -20°C to 120°C temperature measurement range. Camera selection was based on thermal sensitivity, frame rate capability, temperature range, and smartphone integration compatibility. For research consistency, temperature range was constrained to 20-40°C through FLIR mobile application settings, ensuring consistent pixel intensity-to-temperature mapping across all participants.

The ground truth measurement system employs the Fitmate Pro indirect calorimeter, measuring oxygen consumption (VO2) and carbon dioxide production (VCO2) to calculate energy expenditure using established metabolic equations. The device operates on the fundamental principle that energy metabolism involves substrate oxidation in reactions consuming oxygen and producing carbon dioxide in predictable stoichiometric ratios. VO2 conversion to energy expenditure utilizes the formula: kcal/min = (VO2 × body weight × 5 kcal/L) / 1000.

**Signal Processing Pipeline Architecture** begins with region of interest (ROI) identification using classical computer vision techniques applied to parallel RGB video streams. Initial 10 seconds of recording utilizes landmark detection algorithms to precisely locate nostril positions, establishing spatial coordinates guiding thermal signal extraction throughout exercise sessions. This initialization phase is critical because thermal images alone lack spatial detail necessary for automated nostril detection.

The Channel and Spatial Reliability Tracker (CSRT) algorithm maintains nostril tracking throughout exercise sessions, demonstrating superior performance compared to alternatives like MediaPipe and OpenPose for thermal video applications. CSRT effectiveness stems from its dual-channel approach: channel reliability algorithms handle low contrast conditions typical in thermal imagery, while spatial reliability algorithms manage variations in target region temperature and size occurring during exercise.

Respiration signal extraction computes average pixel intensity within tracked nostril ROI according to: I\_avg(t) = (1/WH) × Σ(I\_ROI(x,y,t)), where W and H represent ROI dimensions and I\_ROI denotes pixel intensity at spatial position (x,y) and temporal position t. This averaged intensity signal over time provides direct representation of temperature fluctuations in the nasal cavity caused by respiratory airflow.

**Environmental Control and Calibration Protocols** address critical factors affecting thermal measurements that could compromise system accuracy in real-world deployment scenarios. Data collection occurs in precisely controlled 76°F environments with standardized air circulation patterns to minimize external thermal influences. Participant acclimatization periods prior to data collection ensure environmental temperature effects are minimized and thermal baseline conditions are established.

Temperature feature extraction targets the forehead region, identified through pilot experiments as optimal for detecting exercise-induced temperature changes due to bony prominence characteristics and high vascularization. Physical activity increases metabolic heat generation transmitted through blood vessels and nerves in bony regions, creating measurable skin temperature increases detectable through thermal imaging.

**Machine Learning Architecture and Implementation** implements a novel two-phase Temporal Convolution Network (TCN) approach mimicking physiological processes underlying indirect calorimetry. This innovative architecture was specifically designed to address fundamental challenges of energy expenditure estimation: accounting for individual physiological variations while maintaining generalization across diverse populations and exercise conditions.

Phase one estimates exhaled air volume from respiration rate measurements: v\_t = f1(v\_(t-k:t-1), RR\_(t-k:t)), where v\_t represents estimated volume at time t, k defines temporal context window, and RR represents respiration rate measurements. This phase learns complex relationships between respiration frequency and tidal volume, accounting for individual differences in lung capacity, breathing patterns, and respiratory efficiency.

Phase two predicts VO2 from volume estimates: vo2\_(t+p) = f2(v\_t:v\_(t+p-1)), where p defines volume integration window necessary for accurate VO2 estimation. This phase learns transfer functions between ventilation and oxygen consumption, accounting for individual differences in gas exchange efficiency, dead space ventilation, and metabolic efficiency.

The TCN architecture leverages causal convolutions ensuring unidirectional information flow, preventing data leakage that could artificially improve training performance but fail during deployment. Dilated convolutions capture long-range temporal dependencies crucial for physiological signal analysis, enabling networks to integrate information across multiple breathing cycles and temporal scales.

Hyperparameter optimization through systematic grid search determined optimal network configurations: input chunk size (l=20 corresponding to 60 seconds of data), kernel size (k=5), and dilation base (b=2), resulting in 3-layer networks with 15 kernel filters each. These parameters balance model complexity with computational efficiency while providing sufficient capacity to learn complex physiological relationships.

**Multi-modal Integration Strategies** extend basic respiration-only approaches by incorporating heart rate and temperature data as additional covariates, creating comprehensive physiological monitoring systems. Enhanced model formulation becomes: v\_t = f1(v\_(t-k:t-1), HR\_(t-k:t), T\_(t-k:t), RR\_(t-k:t)), where HR represents heart rate measurements and T denotes temperature measurements.

The multi-modal approach enables systematic evaluation of different sensor combinations, providing insights into relative contributions of different physiological signals and informing optimal sensor fusion strategies for different deployment scenarios. Results demonstrate that while respiration rate alone provides superior performance compared to traditional approaches, integrating multiple physiological signals provides incremental improvements and enhanced robustness.

**Validation Methodology and Statistical Framework** employs Leave-One-Participant-Out Cross-Validation (LOOCV) to ensure robust generalization across individuals while accounting for significant inter-individual physiological variations representing major challenges in energy expenditure estimation. This approach trains models on data from 53 participants while testing on the remaining participant, repeating for all 54 participants to provide unbiased performance estimates.

LOOCV provides critical advantages: maximizes available training data for each fold, provides less biased estimates compared to random train-test splits, and explicitly accounts for individual physiological variations that could introduce systematic errors. The approach is particularly important given wide ranges of individual differences in respiratory patterns, metabolic efficiency, and body composition directly affecting energy expenditure relationships.

**Evaluation and Results**

The comprehensive evaluation of JoulesEye involved 54 carefully selected participants encompassing diverse demographics and fitness levels, designed to provide robust validation across population variations representing major challenges in energy expenditure estimation. The participant cohort included 24 female and 30 male subjects aged 25-54 years, with 41 participants performing cycling on ergometers and 13 conducting treadmill running exercises.

**Participant Demographics and Protocol Design** encompassed diverse body compositions, fitness levels, and age ranges to ensure robust generalization across population variations representing significant challenges in practical energy expenditure estimation. The age range was selected to include both younger adults with typically higher metabolic rates and older adults with typically lower metabolic rates, while excluding populations requiring specialized medical supervision.

The two-session experimental design represents fundamental innovation necessitated by constraints that indirect calorimeter masks completely occlude nostrils, making simultaneous thermal respiration measurement and ground truth energy expenditure collection impossible. Session one focuses on validating thermal respiration rate measurement accuracy by comparing thermal-derived measurements with chest belt references during unmasked exercise. Session two provides ground truth energy expenditure measurements through indirect calorimetry while participants wear face masks.

Exercise protocols were carefully designed to provide diverse energy expenditure patterns while maintaining participant safety and comfort. High-Intensity Interval Training (HIIT) protocols included ergometer resistance varying between maximum and minimum levels every three minutes, or treadmill speeds alternating between 5-6 mph (high intensity) and 2.5-3 mph (low intensity). HIIT approaches ensure diverse energy expenditure patterns, preventing model bias toward gradually increasing metabolic demands typical of steady-state exercise.

**Respiration Rate Accuracy Validation** demonstrates JoulesEye's fundamental capability with exceptional 2.1% MAPE when comparing thermal-derived respiration rates to chest belt references during session one. This performance significantly exceeds previous electrocardiogram and photoplethysmogram-based respiration estimation methods, establishing thermal imaging as superior non-contact respiration monitoring technique during vigorous exercise.

Comparative analysis between indirect calorimeter and chest belt respiration measurements reveals 1.68% MAPE, validating chest belts as acceptable reference standards for bridging two-session experimental design. However, detailed analysis identified specific instances where chest belt measurements failed during vigorous motion, particularly when participants readjusted positions on cycling equipment, highlighting thermal imaging's robustness advantages.

Manual annotation of both thermal and chest belt data provided critical insights into failure modes of different measurement approaches. Visibility of respiration patterns in thermal video enabled identification and correction of chest belt measurement errors, providing more accurate understanding of true respiration rates during exercise. Analysis revealed chest belt failures occurred primarily during sudden movements, equipment adjustments, or when belts shifted position due to perspiration.

**Energy Expenditure Estimation Performance** represents core research contribution, achieving remarkable 5.8% MAPE when using thermal-derived respiration rate compared to indirect calorimetry gold standards. This performance dramatically outperforms Apple Watch estimates (37.6% MAPE), representing more than six-fold improvement and approaching precision levels typically associated with clinical-grade equipment. Even ground truth heart rate data from chest belts yields only 10.2% MAPE, while Apple Watch heart rate produces 12% MAPE.

The comparison with Apple Watch performance provides particularly compelling evidence for JoulesEye's practical value. Apple Watch represents current state-of-the-art in consumer wearable technology, incorporating sophisticated sensors, advanced signal processing algorithms, and machine learning models developed by leading technology companies. JoulesEye achieving more than six-fold better accuracy demonstrates fundamental advantages of thermal-based respiration sensing.

Multi-modal sensor fusion experiments reveal incremental but meaningful improvements: combining estimated respiration rate with Apple Watch heart rate reduces MAPE to 5.5%, while adding temperature data further improves performance to 5.2%. However, these gains remain modest compared to dramatic improvements achieved by incorporating respiration rate itself, emphasizing this physiological signal's fundamental importance.

**Individual Variation Analysis** provides crucial insights into differential performance of energy expenditure estimation approaches across diverse populations. Body composition analysis reveals particularly compelling evidence for respiration rate's physiological advantages: participants with normal BMI show Apple Watch errors of 29.7% versus JoulesEye errors of 5.2%, while overweight participants demonstrate Apple Watch errors of 51.8% compared to JoulesEye errors of 6.9%.

This analysis supports theoretical foundations that respiration rate better captures body composition effects on energy expenditure, explaining why conventional heart rate-only approaches fail dramatically across diverse populations. Differential performance across BMI categories illustrates fundamental limitations of heart rate-based approaches becoming more pronounced in populations with higher adiposity levels.

**Temporal Resolution Analysis** examines practical deployment considerations by systematically varying input window lengths to understand trade-offs between measurement accuracy and response time. While optimal performance requires 90 seconds of data, reducing window length to 30 seconds still achieves 6.4% MAPE, and even 15-second windows achieve 10.1% MAPE compared to 12% for heart rate alone.

The temporal analysis reveals important insights for practical system deployment. The 90-second initialization requirement, while longer than ideal for immediate feedback applications, remains acceptable for many practical scenarios including post-exercise assessment, periodic metabolic monitoring, and integration with continuous heart rate monitoring systems.

**Low-Resolution Feasibility Assessment** addresses critical cost and power consumption concerns through systematic experiments with 32×24 pixel MLX90640 thermal cameras suitable for smartwatch integration. Despite achieving only 3 fps and 8.1% respiration rate MAPE due to motion-induced dithering effects, the low-resolution system produces 15.4% energy expenditure MAPE for respiration alone and 10.1% when combined with heart rate data.

The low-resolution experiments provide crucial proof-of-concept evidence that thermal-based respiration sensing can be implemented in consumer devices at reasonable cost and power consumption levels. The $20 cost estimate for low-resolution thermal sensors makes integration into consumer devices economically viable, particularly given significant accuracy improvements achieved.

**Cross-Activity Performance Validation** demonstrates consistent accuracy across both cycling and running activities, addressing common limitations of previous research focusing on single exercise modalities. While heart rate-based estimates show significantly higher errors for cycling activities, particularly at low intensities where Apple Watch tends to overestimate heart rate responses, respiration rate-based estimates maintain consistent accuracy across both exercise modalities and intensity levels.

The cross-activity validation addresses important limitations in previous energy expenditure research typically focused on single exercise types, potentially creating algorithm bias toward specific movement patterns or metabolic demands. JoulesEye maintaining consistent performance across different exercise modalities suggests robust generalization to diverse physical activities.

**Discussion and Implications**

The exceptional performance achieved by JoulesEye establishes thermal imaging-based respiration sensing as transformative technology for wearable energy expenditure monitoring, with implications extending far beyond fitness applications to encompass clinical medicine, sports science, and public health. The system's 5.8% MAPE compared to indirect calorimetry gold standards represents more than technical achievement—it constitutes a paradigm shift challenging fundamental assumptions about physiological monitoring in consumer devices.

**Physiological Basis for Superior Performance** stems from respiration rate's unique ability to integrate information from multiple physiological systems collectively determining energy expenditure. The neurological connection between brain motor cortex regulation of both respiration rate and aerobic capacity creates direct physiological links absent in heart rate measurements. This explains superior correlation with energy expenditure (0.93 vs 0.78 for heart rate) and captures rapid metabolic changes that cardiovascular responses typically miss due to inherent system delays.

The superior physiological basis extends to respiration rate's natural incorporation of body composition information, particularly adiposity levels directly influencing metabolic efficiency and energy expenditure patterns. Adipose tissue distribution significantly affects respiratory mechanics, lung capacity, and breathing efficiency, creating natural correlations between respiration patterns and metabolic characteristics that heart rate measurements cannot capture.

**Clinical and Medical Applications** present immediate opportunities for transformative impact in healthcare delivery and patient management. The International Society of Sports Nutrition emphasizes that athletes in intensive training may burn over 1200 calories per hour during exercise, requiring precise energy expenditure measurement for optimal performance and injury prevention. Current smartwatch errors exceeding 30% could lead to caloric intake errors of 360+ calories per hour, accumulating to potentially dangerous levels during extended training periods.

JoulesEye's 5.8% error rate approaches precision levels required for clinical applications, potentially enabling medical-grade metabolic monitoring outside traditional clinical settings. This capability could revolutionize chronic disease management by providing physicians and patients with continuous, accurate metabolic data for evidence-based treatment decisions. Diabetes management could benefit enormously from accurate real-time energy expenditure monitoring for insulin dosing decisions, dietary planning, and exercise prescription.

Cardiovascular rehabilitation programs could leverage JoulesEye's accuracy for precise exercise prescription and monitoring, ensuring patients exercise within appropriate intensity ranges for optimal recovery while avoiding dangerous overexertion. Obesity treatment programs could utilize accurate energy expenditure monitoring to provide patients and clinicians with objective data for dietary and exercise interventions.

**Technology Integration and Commercialization Prospects** appear promising despite current engineering challenges, with multiple pathways for practical deployment in consumer devices. The demonstration that 32×24 pixel thermal cameras can achieve reasonable performance (10.1% MAPE when combined with heart rate) provides compelling evidence for smartwatch integration feasibility. At approximately $20 cost for low-resolution thermal sensors, economic barriers for consumer adoption remain manageable.

The temporal requirements—90 seconds for initial estimates with subsequent second-by-second updates—present both challenges and opportunities for practical deployment. While continuous monitoring remains impractical with current technology, periodic respiration assessment combined with continuous heart rate monitoring could provide substantially improved energy expenditure estimates compared to heart rate alone.

**Comparison with Existing Technologies** reveals JoulesEye's transformative potential within current competitive landscapes. Consumer-grade health trackers typically require MAPE ≤10.79% to be considered clinically acceptable according to established validation standards, with many commercial devices showing MAPE >20% in independent validation studies. JoulesEye's 5.2% MAPE approaches the 5% gold standard typically associated with clinical-grade indirect calorimetry.

The multi-modal sensor fusion approach demonstrates clear synergistic benefits suggesting optimal implementation strategies for practical deployment. While respiration rate alone achieves 5.8% MAPE, adding heart rate and temperature data improves performance to 5.2% MAPE. This suggests optimal wearable devices should integrate multiple physiological sensors rather than relying on single modalities.

**Privacy and User Acceptance Considerations** represent important factors for successful deployment, with thermal imaging offering unique advantages in privacy-preserving physiological monitoring. Thermal imaging at low resolution (32×24 pixels) provides physiological information while maintaining complete privacy, as facial features and personal identity information remain completely unidentifiable at such resolutions.

The non-contact nature eliminates concerns about sensor hygiene, skin irritation, and device maintenance that plague contact-based physiological monitoring approaches. Users would not need to maintain skin contact with sensors, clean sensor surfaces, or deal with sensor displacement during exercise, potentially improving long-term adherence to physiological monitoring regimens.

**Broader Impact on Wearable Health Technology** suggests fundamental shifts in how physiological monitoring devices are designed and evaluated. The success of thermal imaging for respiration detection during exercise demonstrates that non-contact physiological monitoring can achieve clinical-grade accuracy in consumer devices, potentially revolutionizing multiple health monitoring applications beyond energy expenditure.

The research methodology employed, particularly systematic comparison with gold-standard clinical measurements and comprehensive validation across diverse populations, establishes new standards for wearable device validation that could improve overall quality and reliability of consumer health technology. The two-session experimental design represents innovative approaches to addressing fundamental constraints in physiological measurement validation.

**Limitations and Future Work**

Despite significant achievements, JoulesEye faces several fundamental limitations constraining immediate deployment and highlighting critical research directions for advancing thermal-based physiological monitoring in consumer devices. These limitations span technical, practical, and methodological domains, each requiring distinct approaches and innovations.

**Real-time Processing Constraints** represent the most significant barrier to practical deployment, with current implementation requiring offline video processing and deep learning pipeline execution after data collection. Computational demands of thermal video processing, region tracking algorithms, and neural network inference significantly exceed current smartwatch processing capabilities, necessitating either cloud-based processing with associated latency and connectivity requirements, or substantial algorithmic optimization for edge deployment.

**Temporal Resolution Limitations** constrain user experience through 90-second initialization periods required for first energy expenditure estimates, proving impractical for immediate fitness feedback applications users expect from wearable devices. While subsequent estimates update every second after initialization, initial delays significantly limit practical utility for applications requiring immediate metabolic assessment.

**Hardware Integration Challenges** persist despite promising low-resolution results, with current prototype thermal cameras achieving only 3 fps and experiencing motion-induced dithering effects degrading performance. Engineering challenges include increasing frame rates while maintaining low power consumption, miniaturizing thermal sensors for practical wearable integration, and developing robust mounting systems maintaining optimal viewing angles during diverse exercise activities.

Future thermal sensor development must address multiple technical challenges simultaneously. Higher frame rate capabilities without proportional power increases require innovations in sensor design, readout electronics, and power management systems. Improved sensitivity enabling operation at ultra-low resolutions could reduce data processing requirements while maintaining physiological signal quality.

**Environmental Sensitivity and Robustness** remains a concern despite controlled laboratory validation, with real-world deployment requiring operation across diverse environmental conditions including varying ambient temperatures, humidity levels, air circulation patterns, and other factors potentially affecting thermal measurements. Environmental factors could affect thermal measurements through multiple mechanisms requiring systematic characterization and potentially algorithmic compensation.

Future work must include comprehensive environmental testing across ranges of conditions users might encounter during exercise activities, including indoor and outdoor environments, different climate conditions, various exercise equipment and facilities, and integration with other wearable devices that might introduce interference.

**Conclusion**

JoulesEye is a groundbreaking wearable technology that uses thermal imaging for respiration-based energy expenditure monitoring, achieving clinical-grade accuracy (5.8% MAPE) far surpassing current commercial smartwatches. It proves thermal imaging as a viable, cost-effective, and accurate approach for consumer devices, with broad applicability across different populations and exercise types. The technology has transformative potential beyond fitness, including clinical medicine, sports science, and public health, enabling continuous metabolic monitoring and personalized health interventions. Combining multiple sensors (respiration, heart rate, temperature) further improves accuracy (5.2% MAPE), highlighting the benefits of multi-modal wearables. While challenges remain in real-time processing and integration, JoulesEye lays a strong foundation for future advancements in wearable health monitoring.

**PROPOSAL**

**Research Proposal**

**Abstract**

Energy expenditure (EE) measurement is crucial for health monitoring, weight management, and athletic performance optimization. Current consumer wearable devices often rely predominantly on heart rate and motion sensors, which have large errors in calorie estimation, especially during vigorous exercise. The recently proposed JouleEye system demonstrates that respiration rate estimates derived from thermal imaging can significantly improve EE accuracy, outperforming popular wrist-worn devices. Building on this promising technology, this proposal outlines a comprehensive 1-year research plan to evolve JouleEye from a proof-of-concept into a practical, real-time, wrist-worn system. The research aims to develop miniaturized thermal imaging hardware, real-time onboard processing, optimized respiratory signal tracking under motion and occlusion, and enhanced machine learning models integrating multimodal inputs to accurately estimate EE. Data will be collected from a diverse population across multiple exercise modalities. The outcome promises a next-generation wearable capable of accurate, privacy-respecting, and user-friendly EE monitoring for everyday fitness and clinical applications.

**1. Introduction and Motivation**

* 1. **Importance of Energy Expenditure**

The accurate measurement of energy expenditure (EE) is fundamental to a wide range of health, fitness, and medical applications. EE quantifies the number of calories burned per unit time and serves as a proxy for physiological effort and metabolic activity during physical tasks. This metric is crucial for: Weight management, by enabling precise balancing of caloric intake and expenditure to support healthy body composition. Athletic training, where monitoring EE helps tailor workouts for optimized performance and recovery. Metabolic disorder diagnosis, as deviations in EE patterns may indicate underlying conditions such as thyroid dysfunction or respiratory impairments. Rehabilitation monitoring, facilitating adjustment of therapy intensity in recovery from injury or illness.

**1.2 Limitations of Current Consumer Wearables**

Despite widespread adoption, existing consumer-grade wearables predominantly estimate EE using heart rate (HR) and accelerometry data, which are subject to notable limitations:

* High estimation errors: Studies have demonstrated EE inaccuracies exceeding 40% in typical consumer devices, attributable to factors including sensor noise, motion artifacts, and oversimplified predictive models that overlook personal physiological variation.
* Poor robustness during specific activities: For example, in cycling where hand movement is minimal, HR-based EE algorithms can underperform due to reduced cardiovascular signal dynamics.
* Neglect of intrinsic physiological differences: Variations in body composition, fitness levels, and metabolic efficiency are typically not incorporated into existing estimators, limiting their generalizability and personalization.
* Inherent limitations of accelerometry: Movement data alone does not capture internal metabolic changes or body state, especially during activities lacking distinctive motion patterns.

As a result, current wearables provide suboptimal insights into actual physiological energy use, restricting their utility for precise health management or clinical applications.

**1.3 JouleEye: A Promising Alternative**

The JouleEye system innovatively leverages thermal infrared imaging to measure respiration patterns by capturing periodic temperature fluctuations near the nostrils caused by airflow during breathing. Respiration rate (RR) serves as a more direct marker of metabolic effort and correlates with body composition—factors that strongly influence EE.

Key demonstrated achievements of JouleEye include:

* A significantly lower mean absolute percentage error (MAPE) of 5.8% in EE estimation across 54 participants, compared to roughly 37% MAPE reported for Apple Watch estimates based on HR alone.
* Robust respiration sensing during vigorous motion such as running and cycling, overcoming common challenges faced by other sensing modalities in dynamic environments.
* Demonstrated feasibility of using low-cost, low-resolution thermal sensors compatible with wearable form factors, enabling potential integration into smartwatches or fitness bands.

Moreover, combining respiration data with HR and facial temperature estimates further enhances EE estimation accuracy, highlighting the added value of multimodal sensing.

1.4 Research Gap and Opportunity

While JouleEye exhibits compelling accuracy improvements, its current prototype implementation relies on a smartphone-attached thermal camera and offline processing pipelines, incurring practical limitations:

* A relatively long data acquisition window of approximately 90 seconds is required to generate reliable EE estimates, posing usability constraints for real-time feedback.
* Integration challenges remain for embedding thermal sensing hardware in compact wearable devices such as smartwatches with real-time on-device computation.
* Addressing motion artifacts, occlusion scenarios, and environmental temperature variability in unconstrained daily settings is necessary to ensure robustness.

This research opens opportunities to innovate wearable thermal sensing platforms that deliver near real-time EE monitoring, preserve privacy by avoiding RGB video, and scale cost-effectively. Advancing JouleEye towards fully integrated, responsive wearable solutions can transform personalized fitness tracking, clinical monitoring, and health management.

**2. Objectives**

The key objectives over the next 1-year period are designed to advance JoulesEye from a research prototype towards a practical, real-world wearable solution for accurate energy expenditure (EE) estimation:

* **Develop a miniaturized, wearable thermal imaging module**This includes designing and engineering a compact thermal camera system with sufficient spatial resolution and frame rate that can be seamlessly integrated into a smartwatch form factor. The module must maintain power efficiency and thermal sensitivity to reliably detect subtle temperature fluctuations around the nostrils during various physical activities.
* **Implement real-time respiratory signal extraction and tracking onboard**Develop robust algorithms capable of realtime processing on embedded hardware within the wearable device. These methods will extract and continuously track the respiration signal from thermal video, ensuring resilience to vigorous motion (running, cycling) and frequent occlusions (e.g., hand movements covering the face). This will replace offline processing to enable immediate feedback to the wearer.
* **Optimize deep learning models for multi-modal EE estimation**Refine and compress Temporal Convolutional Network (TCN) architectures or alternative lightweight models to accurately estimate energy expenditure by fusing respiration rate (RR), heart rate (HR), and facial temperature signals. The models will be trained and validated over diverse exercise modalities to maximize accuracy and generalizability across populations.
* **Reduce user look time to under 30 seconds**  
  Shorten the duration users need to glance at or interact with the device to obtain a reliable EE estimate. This improvement aims to enhance usability and acceptability by minimizing disruption during exercise, overcoming the current limitation of requiring approximately 90 seconds of continuous data.
* **Design and execute a comprehensive data collection protocol**Conduct a systematic study involving a demographically and physiologically diverse group of participants engaging in varied physical activities. This data acquisition will support model training, evaluation of robustness across different conditions, and validation against gold standard indirect calorimetry.
* **Incorporate uncertainty quantification for EE estimates**  
  Integrate methods to estimate and communicate confidence intervals or uncertainty levels associated with EE predictions. This is crucial for real-world deployments where sensor noise, environmental factors, or user behavior may impact measurement reliability and should be transparently conveyed.
* **Conduct preliminary usability and acceptability studies**  
  Evaluate the smartwatch prototype and interaction paradigms with target users to understand practical challenges, comfort, user experience, and adoption barriers. Feedback will inform iterative hardware and software design improvements to facilitate seamless integration into daily fitness monitoring routines.

**3. Proposed Methodology**

**3.1 Hardware Development**

To realize a wearable energy expenditure estimation system based on respiration sensing, the hardware development will focus on miniaturization, multimodal sensing integration, and seamless user interaction. Key components and design considerations include:

**Selection of Thermal Sensor**

* The core sensing unit will leverage a compact thermal infrared sensor capable of capturing subtle temperature variations caused by breathing patterns near the nostrils. A suitable candidate is the MLX90640 thermal array sensor, which balances size, power consumption, and performance.
* **Key specifications for selection include:**
  + Frame rate of at least 8 frames per second (FPS) to reliably track dynamic respiratory cycles during vigorous motion.
  + Spatial resolution of approximately 32×24 pixels or higher, sufficient to localize and resolve temperature fluctuations at the nostril region while maintaining low component cost and power requirements.
  + Sensitivity across the relevant temperature range (approximately human skin temperature variation) to detect subtle breathing-induced thermal changes.

**Wearable Integration**

* A custom printed circuit board (PCB) will be designed to host the thermal sensor, supporting components, and necessary power management circuits.
* The PCB will be engineered to fit within the ergonomic constraints of a wrist-worn form factor, considering factors such as size, weight, heat dissipation, and user comfort.
* A lightweight, durable enclosure will encapsulate the PCB and sensor element, incorporating protective yet thermally transparent materials that do not obstruct infrared sensing.
* Mechanical design will enable secure attachment to the user’s wrist, with consideration for adjustability to accommodate different wrist sizes.

**Additional Sensors**

* An auxiliary low-resolution RGB camera will be integrated adjacent to the thermal sensor to provide visual context for initial nostril detection. This is critical for robust region-of-interest localization before autonomous tracking takes over, especially in cases of partial occlusion or motion-induced displacement.
* Photoplethysmogram (PPG) sensors embedded within the wearable will measure heart rate concurrently, enabling multimodal physiological data fusion to enhance the accuracy of energy expenditure estimation.
* Sensor placement and optical transparency will be carefully optimized to ensure minimal interference among sensing modalities.

**Legacy Smartphone Interface**

* A communications interface (e.g., Bluetooth Low Energy) will facilitate reliable data transfer from the wearable module to a companion smartphone or device for further processing, data storage, and user feedback display.
* The smartphone application will present an intuitive user interface (UI) allowing users to initiate measurements, monitor live physiological signals, and review energy expenditure estimates.
* Firmware and app development will emphasize efficient data synchronization, battery management, and privacy-preserving data handling, ensuring scalability and user trust.

By combining these hardware components and design principles, the project aims to develop a fully integrated, compact, and user-friendly wearable system capable of capturing high-quality thermal and physiological data necessary for real-time, accurate energy expenditure estimation.

**3.2 Real-Time Respiratory Signal Processing**

To achieve accurate and continuous respiratory monitoring during exercise, especially under motion and occlusion, the following multi-stage processing pipeline is implemented:

**Automatic ROI Detection**

* Leverage the RGB camera frames, which provide higher-resolution and visually richer data, to perform automatic detection of facial landmarks, specifically localizing the nostrils.
* Utilize lightweight and efficient facial landmark detection frameworks, such as MediaPipe or comparable algorithms, capable of real-time operation on embedded or wearable hardware.
* This initial nostril localization step defines the Region of Interest (ROI) on the face where respiratory-induced thermal variations are expected — typically the vicinity around the nostrils.
* The robust detection in RGB is essential for initializing the pipeline and recalibrating ROI when drifting or tracking loss occurs in the thermal image.

**ROI Tracking in Infrared (IR) Frames**

* After initial nostril region detection, track the ROI over the sequence of thermal frames where RGB data may be less reliable, unavailable, or privacy-constrained.
* Employ an optimized Channel and Spatial Reliability Tracker (CSRT) algorithm tailored for thermal imaging, which uses:
  + Channel reliability: to weigh contributions from different spectral channels in the IR frames, enhancing feature stability despite low contrast or sensor noise.
  + Spatial reliability: to handle shape or scale variations and partial occlusions by dynamically adjusting the region confidence.
* This tracking approach maintains the precise localization of the nostril region even under vigorous body movements, changes in head orientation, and momentary occlusions such as hand gestures or sweat wiping.

**Respiratory Signal Extraction**

* Compute the mean pixel intensity within the ROI for each infrared frame, as the thermally-sensitive brightness fluctuations correspond to temperature changes caused by airflow during inhalation and exhalation.
* Apply temporal signal processing techniques to the raw intensity time series, including Gaussian smoothing filters, which reduce high-frequency noise and enhance the respiratory waveform’s clarity.
* Interpolate the smoothed signal to produce a uniformly sampled signal stream, compensating for any frame drops or non-uniform frame rates, thereby ensuring consistent data input for downstream respiration rate estimation models.

**Resilience to Occlusion and Tracking Failures**

* Implement a fallback mechanism whereby if the thermal ROI tracking consistently fails or deviates beyond a threshold—due to prolonged occlusion or rapid motion—the system automatically triggers the RGB-based nostril localization to reacquire and reset the ROI.
* This multi-modal redundancy ensures continuous respiratory signal availability, mitigating data loss impacts and maintaining accuracy in real-time usage scenarios.

By combining advanced, optimized tracking algorithms with cross-modal sensing and adaptive recovery strategies, the respiratory signal processing pipeline provides reliable and robust extraction of respiration dynamics critical for accurate energy expenditure estimation during real-world exercise.

**3.3 Multimodal Machine Learning Pipeline**

**Input Data Fusion**A comprehensive physiological input vector is constructed by combining multiple complementary data streams:

* **Respiration Rate (RR):** Extracted from thermal infrared (IR) imaging focused on the nostril region, this captures the breathing pattern by detecting temperature changes caused by airflow during inhalation and exhalation. RR serves as a direct indicator of respiratory effort and metabolic demand.
* **Heart Rate (HR):** Derived from photoplethysmography (PPG) sensors, typically integrated in wearable devices such as smartwatches. HR provides cardiovascular context reflecting the circulatory response to exercise intensity.
* **Facial Temperature:** Obtained from thermal IR images of facial regions with bony prominences (e.g., forehead), which exhibit temperature elevations correlating with metabolic heat production and thermoregulatory blood flow changes during exertion.  
  These synchronized multimodal signals yield a rich feature set that collectively represents the body's physiological state during physical activity, improving the accuracy and robustness of energy expenditure (EE) estimation beyond single-modality approaches.

**Model Architecture**The pipeline employs a two-stage cascaded deep learning framework based on Temporal Convolutional Networks (TCNs) designed for sequence modeling and time series forecasting with the following specifications:

* **Stage 1 -** Volume Estimation Network: This TCN receives the time series of RR along with optionally HR and facial temperature covariates and predicts the volume of exhaled air over time. This intermediate output approximates pulmonary ventilation, a key physiological determinant influencing oxygen consumption. The model leverages causal dilated convolutions to capture temporal dependencies without information leakage, allowing long-range context modeling essential for breathing dynamics.
* **Stage 2 -** VO2 Prediction Network: Feeding on the predicted exhaled volume from Stage 1, this second TCN estimates the oxygen consumption (VO2) continuously. VO2 directly scales with energy expenditure and serves as the final model output. This cascaded approach modularizes the physiological inference tasks and simplifies learning complex, non-linear mappings from raw sensor data to EE.  
  Model residuals and hyperparameters such as kernel size, dilation factors, and layer depths are optimized to balance accuracy and computational efficiency for embedded deployment.

**Data Augmentation and Regularization**To enhance generalization and robustness, the training process incorporates noise injection and augmentation strategies:

* Synthetic noise is applied to input signals during training to simulate various real-world sensor disturbances including motion artifacts, signal dropouts, and environmental noise.
* This approach prevents overfitting and ensures model resilience when deployed in uncontrolled exercise scenarios where sensors face variable conditions.
* Additional regularization techniques such as dropout and early stopping are employed to stabilize training and avoid overfitting given the limited size and diversity of physiological datasets.

**Uncertainty Estimation**Recognizing the inherent variability and potential inaccuracies in physiological sensing, the pipeline integrates uncertainty quantification into EE predictions:

* **Bayesian Dropout:** Dropout layers remain active during inference, enabling multiple stochastic forward passes. The variance across these passes provides an empirical measure of prediction uncertainty, which can guide reliability assessment and downstream decision-making.
* **Model Ensembles:** An alternative approach employs multiple independently trained models whose outputs are aggregated. Prediction confidence intervals are derived from ensemble variance, strengthening robustness against individual model biases.  
  Estimating uncertainty helps inform users or clinical applications about the confidence of calorie burn estimates, supporting safer and more informed usage.

**Model Deployment Optimization**To realize real-time energy expenditure estimation on resource-constrained wearable platforms, several model compression and acceleration techniques are applied:

* **Quantization:** Reducing model parameter precision from floating-point to fixed-point or integer formats shrinks memory footprint and enables faster arithmetic operations on embedded CPUs or microcontrollers.
* **Pruning:** Removing redundant or less influential weights and filters decreases model size and computational load without significant loss of accuracy.
* **Hardware-aware Optimization:** The models are tuned for specific embedded platforms to exploit available instructions, cache sizes, and parallelism for maximum efficiency.These optimizations collectively enable near real-time, continuous inference of EE on wearable devices with limited processing and power resources, promoting user convenience and battery life.

This multimodal machine learning pipeline synergizes complementary physiological data and state-of-the-art deep learning architectures with practical deployment considerations to achieve accurate, robust, and timely energy expenditure estimation during exercise. If needed, additional details on training regimes, software frameworks, or sensor synchronization can be provided.

**3.4 Data Collection and Validation**

**Participant Recruitment**The study targets recruiting a diverse cohort of approximately 60 to 80 healthy adult volunteers spanning:

* Age Range: 18 to 60 years, ensuring representation across early adulthood to middle age.
* Body Mass Index (BMI) Categories: Including normal weight, overweight, and obese individuals to assess model performance variability with body composition.
* Gender Balance: Aim for near-equal distribution of male and female participants to account for physiological differences.  
  This diversity supports generalizability of the energy expenditure estimation model across different demographics.

**Exercise Modalities**Participants engage in a range of physical activities representative of common exercise:

* Treadmill Running: Continuous locomotion at varying speeds and intensities.
* Cycling: Stationary bicycle ergometer sessions incorporating various resistance settings.
* Walking: Controlled walking at different paces to capture low-impact activity respiration patterns.
* Stair Climbing: Simulated or actual stair ascent and descent to include vertical movement stress.
* Resistance Training: Incorporate bodyweight or weighted exercises to test performance under intermittent effort and upper-limb motion.  
  A broad activity portfolio ensures robustness of the model to different movement patterns and metabolic demands.

**Ground Truth Measurements**

* Use a medical-grade indirect calorimeter to continuously measure oxygen consumption (VO2) and carbon dioxide output during exercise sessions.
* The indirect calorimeter serves as the gold standard reference for energy expenditure, providing accurate EE values for model training and validation.
* Metabolic data are synchronized with wearable sensor streams and video recordings for precise temporal alignment.

**Reference Sensors and Comparisons**

* Employ a chest-worn respiration belt to capture respiratory effort as an auxiliary reference for respiratory rate accuracy.
* Use commercial smartwatches to obtain heart rate and accelerometry data, enabling side-by-side benchmarking of EE estimation performance against existing consumer technology.  
  These references provide comparative baselines and facilitate validation of the novel thermal imaging approach.

**Protocol Design**

* Design exercise sessions to include both continuous steady-state activity (e.g., sustained running or cycling) and intermittent variable-intensity protocols (e.g., interval training, resistance sets).
* Incorporate periods of both smooth and abrupt motion, simulating realistic user behavior including rest breaks, accelerations, and changes in effort level.
* Standardize session duration and environmental conditions (e.g., indoor lab with controlled temperature) to reduce confounding factors.

**Ethical Compliance**

* The study protocol undergoes rigorous review and approval by an institutional review board (IRB) or ethics committee ensuring participant safety and data privacy.
* Participants provide informed consent prior to data collection, acknowledging the study’s scope, potential risks, and use of collected data.
* Follow strict guidelines on data security and anonymization to comply with applicable regulations such as GDPR or HIPAA.

**3.5 Evaluation Metrics**

**Energy Expenditure (EE) Estimation Accuracy**

* The primary metric for assessing model performance is the Mean Absolute Percentage Error (MAPE) between the estimated energy expenditure and the gold-standard ground truth values obtained via indirect calorimetry.
* MAPE quantifies the average absolute difference percentage, reflecting how closely the model predicts true calorie burn rates during various exercise intensities and modalities.
* Lower MAPE values indicate higher estimation fidelity, with thresholds aligned to clinical and consumer-grade device standards for acceptance.

**Respiration Rate (RR) Accuracy**

* The respiration rate estimated from thermal imaging is validated against reference measurements from chest-worn respiration belts considered reliable under controlled conditions.
* Additionally, manual annotation of respiratory cycles from thermal video frames provides an independent ground truth benchmark, particularly useful during motion phases where sensor noise is prevalent.
* Accuracy is summarized using metrics such as mean absolute error, correlation coefficients, and percent error, demonstrating robustness of the signal extraction pipeline amid motion artifacts.

**Latency and Throughput**

* System responsiveness is assessed by measuring the end-to-end latency, defined as the elapsed time from initial data capture (thermal frame acquisition) to the final energy expenditure output from the predictive model.
* Throughput metrics capture the number of inference operations or EE estimates generated per second, crucial for real-time or near-real-time application feasibility.
* Performance tests involve running the full pipeline on the target hardware (e.g., smartphone or wearable MCU) to identify bottlenecks and optimize processing time.

**User Acceptability and Experience**

* Participants complete structured surveys post-experiment to evaluate subjective factors including:
  + Comfort of wearing sensors and interacting with the device.
  + Usability of device interfaces for initiating recordings and viewing outputs.
  + Satisfaction regarding perceived accuracy, convenience, and overall user experience.
* Qualitative feedback helps identify barriers to practical deployment and informs iterative design improvements for consumer adoption.

Together, these comprehensive evaluation criteria ensure the developed system delivers accurate, timely, and user-friendly energy expenditure estimation suitable for realistic exercise monitoring scenarios.

**4. Expected Outcomes and Contributions**

* **Prototype Wearable JouleEye Device**Develop a compact, wrist-worn JouleEye prototype integrating a low-resolution thermal camera capable of delivering high-accuracy energy expenditure (EE) estimation during varied physical activities. The device emphasizes form factor suitability for daily wear and convenience without compromising measurement fidelity.
* **Robust Real-Time Respiratory Signal Tracking**Demonstrate a real-time respiratory rate tracking system that accurately captures respiration signals despite vigorous user motion and common occlusions (e.g., hand movements or temporary loss of sight of nostrils). This system exhibits robustness needed for practical deployment in dynamic exercise environments.
* **Reduced User Interaction Time**Achieve the reduction of user ‘look-time’—the duration during which the user must maintain the device’s line of sight with their face—for obtaining an EE estimate, targeting under 30 seconds while maintaining modeling accuracy. This improvement addresses usability and convenience barriers typical of continuous physiological monitoring.
* **Advancement via Multimodal Fusion Modeling**Provide empirical validation of multimodal sensor fusion models that jointly leverage respiration rate (RR), heart rate (HR), and facial temperature inputs. These models outperform existing commercial devices reliant solely on heart rate, delivering significantly lower EE estimation errors and increased reliability across varying exercise modalities and participant profiles.
* **Framework for Uncertainty-Aware Estimation**Develop an analytical framework incorporating uncertainty quantification within EE estimation, enabling confidence-aware predictions. This approach improves robustness to variable sensing conditions (e.g., different lighting, motion artifacts) and supports more informed decision-making in health and fitness applications.
* **Open-Source Multimodal Dataset**Release a comprehensive, synchronized dataset combining thermal and RGB video, respiration signals, heart rate data, and indirect calorimetry-based oxygen consumption (VO2) measurements. Collected from a demographically diverse cohort, the dataset supports reproducibility, facilitates benchmarking, and accelerates research on non-invasive physiological monitoring and energy expenditure estimation.

Together, these outcomes advance wearable health technology by bridging accuracy, practicality, and data transparency, paving the way for widespread adoption of respiration-informed EE monitoring in consumer devices.

**5. Conclusion**

This research aims to advance the promising JouleEye technology from its experimental stage towards a practical, user-friendly wearable system capable of accurately monitoring energy expenditure in real-time using respiration signals. By integrating low-cost thermal imaging sensors with robust tracking and machine learning models, JouleEye addresses key limitations of existing wearables that rely predominantly on heart rate data. The success of this technology will enable more precise and reliable fitness and health tracking for millions of users worldwide, supporting personalized exercise management and clinical assessments. This work paves the way for novel, non-invasive methods that can transform how energy expenditure is measured both in everyday fitness applications and potentially in medical contexts.