



Multimodal Mental Health Screening from Student Behavioral and Physical-Activity Indicators

Dennis Wanjiru, Mercy Omondi

*Department of Psychiatry, School of Medicine, College of Health Sciences, University of Nairobi, Nairobi
P.O. Box 30197-00100, Kenya*

Abstract

The increasing prevalence of mental health disorders among university students has become a pressing global public health concern. Traditional methods for assessing student psychological wellbeing rely predominantly on self-reported questionnaires and clinical interviews, which are often limited by recall bias, low response rates, and the stigma associated with seeking mental health support. This paper explores the development and evaluation of a comprehensive multimodal mental health screening framework that utilizes passive data collected from everyday student behaviors and physical activity indicators. By continuously monitoring digital phenotypes through smartphones and wearable fitness trackers, an unprecedented level of objective, longitudinal data can be harvested. Behavioral indicators include smartphone usage patterns, mobility traces, and digital social interactions, whereas physical activity indicators encompass step counts, heart rate variability, and sleep duration. The proposed framework aligns and fuses these distinct modalities to predict the onset and severity of stress, anxiety, and depression. Utilizing advanced machine learning techniques, specifically ensemble methods and deep neural network architectures, the system extracts critical temporal and spatial features to model the complex underlying relationships between physical routines, digital habits, and psychological states. This research demonstrates that a multimodal approach significantly outperforms unimodal baselines, providing a scalable, non-intrusive, and highly accurate early warning system for educational institutions. The findings emphasize the necessity of integrating diverse data streams to capture a holistic view of student wellbeing, thereby enabling timely interventions and personalized support mechanisms.

Keywords: *Multimodal Screening, Mental Health, Wearable Sensors, Student Wellbeing*

1. Introduction

1.1 Context and Motivation

The mental health of university students is a subject of growing alarm for educational institutions, healthcare providers, and policymakers worldwide. The transition to higher education represents a critical developmental period characterized by significant academic pressure, financial stress, changes in social support networks, and newfound independence. These concurrent stressors frequently precipitate psychological distress, making this demographic particularly vulnerable to the onset of severe mental health conditions such as generalized anxiety disorder, major depressive disorder, and chronic stress [1]. Despite the clear need for intervention, a substantial proportion of students suffering from psychological distress do not receive adequate or timely clinical care. This treatment gap is exacerbated by institutional resource constraints, the social stigma surrounding psychiatric diagnoses, and the episodic nature of traditional screening programs. Campus health centers conventionally deploy screening mechanisms that depend almost exclusively on subjective self-report instruments administered during infrequent clinical encounters.

These instruments suffer from critical limitations, including social desirability bias, memory distortions, and the inherent inability to capture the dynamic, day-to-day fluctuations in a student's psychological state [2]. Consequently, there is an urgent need to develop objective, continuous, and ecologically valid methodologies for monitoring student mental health. The widespread adoption of consumer electronics, particularly smartphones and wearable fitness trackers, presents a transformative opportunity to address this monitoring gap. These ubiquitous devices are equipped with a diverse array of sensors capable of passively and continuously recording an individual's behavioral and physiological data in real-world settings [3]. This continuous data stream, often referred to as a digital phenotype, provides a granular, longitudinal record of a student's daily life without requiring active user input. By analyzing how students interact with their digital devices and monitoring their physical activity levels, researchers can identify subtle deviations from baseline patterns that may serve as early indicators of mental health deterioration. However, translating raw sensor data into clinically meaningful insights is a profoundly complex endeavor that requires sophisticated data processing and modeling strategies [4].

1.2 Problem Statement and Objectives

While individual modalities, such as physical activity tracking or smartphone usage logging, have shown promise in predicting mental health outcomes, they often provide an incomplete picture of an individual's overall wellbeing. Human behavior and psychological states are multidimensional constructs; therefore, relying on a single data source can lead to high rates of false positives and a lack of contextual understanding. For instance, a sudden decrease in physical activity could be indicative of a depressive episode, but it could equally result from a physical illness or an intense period of academic studying. To disambiguate such scenarios and construct a robust screening tool, it is imperative to integrate multiple, complementary streams of data. The primary objective of this paper is to investigate the efficacy of a multimodal mental health screening system that synergistically combines behavioral indicators extracted from smartphone usage with physical activity metrics derived from wearable sensors. By systematically fusing these distinct modalities, this research aims to construct a highly accurate, generalized model for detecting elevated levels of stress, anxiety, and depression among university students. The study will thoroughly detail the data acquisition architecture, the rigorous preprocessing and feature extraction methodologies required to handle noisy, asynchronous sensor data, and the evaluation of various machine learning algorithms optimized for multimodal fusion. Ultimately, this research seeks to validate the hypothesis that a multimodal approach yields superior predictive performance compared to isolated unimodal models, thereby paving the way for proactive, scalable, and non-intrusive mental health interventions in educational environments.

2. Theoretical Background and Literature Review

2.1 Traditional Approaches to Mental Health Assessment

Historically, the evaluation of psychological wellbeing in academic populations has been dominated by standardized psychometric scales. Instruments such as the Patient Health Questionnaire for depression, the Generalized Anxiety Disorder scale, and the Perceived Stress Scale have established themselves as the gold standard in both clinical and research settings [5]. These tools are highly validated, relatively easy to administer, and provide quantifiable metrics that align with established diagnostic criteria. However, their utility as continuous screening mechanisms is inherently limited by their cross-sectional nature [6]. When students complete these questionnaires, they are essentially providing a retroactive summary of their feelings over a specific preceding period, typically the past two weeks. This retrospective reporting introduces significant recall bias, as individuals experiencing current distress may disproportionately remember negative events while overlooking neutral or positive occurrences [7]. Furthermore, the administration of these surveys demands active participation, which often leads to survey fatigue and declining response rates over time. In a bustling university environment, students are unlikely to consistently engage with daily or weekly self-assessments, rendering longitudinal tracking nearly impossible [8]. The reliance on self-reporting also assumes that students possess the requisite self-awareness to accurately identify and articulate their symptoms, which is not always the case, particularly during the initial stages of a psychological crisis [9]. As educational institutions increasingly recognize the necessity of early intervention, the limitations of point-in-time, subjective assessments have catalyzed a

paradigm shift toward passive, continuous monitoring technologies [10].

2.2 Unimodal Digital Phenotyping

The concept of digital phenotyping, defined as the moment-by-moment quantification of the individual-level human phenotype in situ using data from personal digital devices, emerged as a compelling solution to the constraints of traditional assessments. Early investigations in this domain predominantly focused on unimodal approaches, leveraging either smartphone telemetry or wearable sensor data in isolation. Smartphone-based studies have extensively explored behavioral markers of mental health [11]. Metrics such as screen unlock frequency, total screen time, and the duration of specific application usage have been correlated with varying degrees of psychological distress. For example, excessive engagement with social media applications, coupled with irregular sleep-wake cycles inferred from late-night device interaction, has frequently been associated with elevated depressive symptoms [12]. Additionally, location data acquired via global positioning systems can reveal a user's mobility patterns, wherein decreased spatial exploration and prolonged periods spent in residential locations serve as proxies for social withdrawal and lethargy [13]. Conversely, research utilizing wearable devices has concentrated on physiological and physical activity indicators. Fitness trackers and smartwatches equipped with accelerometers and photoplethysmography sensors enable the continuous monitoring of step counts, energy expenditure, and cardiac parameters. Reductions in daily physical activity and disruptions in sleep architecture are classic somatic symptoms of various mood disorders [14]. Furthermore, metrics derived from heart rate variability offer insights into autonomic nervous system regulation, providing a physiological window into an individual's acute and chronic stress levels [15]. Despite the valuable insights generated by these unimodal studies, their predictive accuracy often plateaus. Environmental confounders, technical anomalies, and the multifaceted nature of mental illnesses dictate that single-stream data cannot fully capture the nuanced reality of a student's lived experience [16].

2.3 The Shift Toward Multimodal Fusion

Recognizing the bottlenecks associated with single-sensor methodologies, contemporary research has decisively pivoted toward multimodal frameworks. By aggregating heterogeneous data sources, researchers can triangulate an individual's behavioral and physiological state, thereby minimizing the impact of noise and ambiguous signals inherent in any single modality [17]. A multimodal approach allows for the contextualization of behaviors; for instance, interpreting a low step count in conjunction with prolonged academic app usage provides a vastly different clinical picture than observing low physical activity alongside high gaming app usage and disrupted sleep [18]. The transition to multimodal analysis introduces complex computational challenges, primarily concerning data synchronization, feature alignment, and the selection of appropriate fusion strategies. Early fusion techniques attempt to concatenate raw or minimally processed data from all sensors before feeding the combined matrix into a predictive model [19]. While conceptually straightforward, this method struggles with the curse of dimensionality and the disparate sampling rates characteristic of different devices. Late fusion, alternatively, trains independent models for each modality and subsequently aggregates their predictions, typically yielding more robust performance by allowing modality-specific algorithms to optimize local feature extraction [20]. Recent advancements have also explored hybrid and intermediate fusion architectures, heavily relying on deep learning networks to autonomously discover latent cross-modal representations [21]. This paper builds upon these foundational concepts, proposing a specialized multimodal pipeline optimized for the unique behavioral and physical markers of university students.

3. System Architecture and Multimodal Data Acquisition

3.1 Overall Framework Design

Developing a reliable mental health screening system requires a robust and secure technological infrastructure capable of seamlessly collecting, transmitting, and storing vast quantities of sensitive personal data. The proposed system architecture is designed with a strong emphasis on unobtrusiveness, ensuring that the data collection process does not interfere with the students' daily routines or introduce observation bias. The architecture comprises three primary tiers: the data acquisition tier located on the user's personal devices, the secure data transmission and storage tier hosted on a centralized cloud server, and the analytics tier where preprocessing and machine learning operations are executed [22].

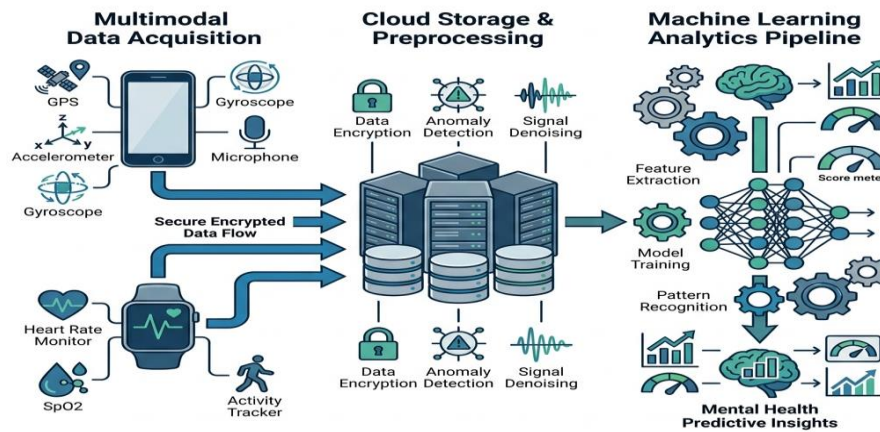


Figure 1 Comprehensive System Architecture Diagram

The data acquisition tier utilizes a custom-developed background application installed on the students' primary smartphones, paired via low-energy wireless protocols to a standardized wrist-worn fitness tracker. This continuous pairing allows the smartphone application to serve as a localized hub, periodically querying the wearable device for accumulated physiological data while simultaneously logging its own telemetry. To mitigate privacy concerns and encourage participation, all collected data is subjected to strict anonymization protocols at the point of origin [23]. Identifiable information, such as specific geographic coordinates or granular communication content, is transformed into abstract behavioral metrics before ever leaving the user's device.

3.2 Sensing Modalities and Variables

The system meticulously gathers an extensive array of variables categorized into two distinct modalities: behavioral indicators and physical activity indicators. Behavioral indicators are predominantly harvested from smartphone interaction logs. Key variables include screen state transitions, which calculate the frequency and duration of device usage sessions throughout the day and night. Application usage categorization is another critical metric; the system logs the aggregate time spent on distinct application classes such as social media, entertainment, productivity, and communication, providing insights into digital coping mechanisms and social connectivity [24]. Furthermore, ambient light sensors and screen activation patterns during typical sleeping hours are utilized to estimate digital sleep disruption, a potent behavioral marker of anxiety and insomnia [25]. Simultaneously, the physical activity modality is populated with data streamed from the wearable fitness tracker. The primary variables include continuous heart rate monitoring, from which both resting heart rate and diurnal heart rate variability are derived [26]. The built-in triaxial accelerometer provides highly granular estimations of step counts, classifying periods of the day into sedentary, lightly active, and vigorously active states. Crucially, the wearable device offers detailed sleep architecture analysis, generating metrics on total sleep time, sleep onset latency, and the frequency of nocturnal awakenings. By capturing these physiological and kinetic parameters, the system establishes a comprehensive baseline of physical wellbeing, essential for identifying the somatic manifestations of psychological distress [27].

4. Data Preprocessing and Feature Extraction Methodology

4.1 Data Cleaning and Synchronization

The raw data harvested from smartphones and wearable sensors in uncontrolled, free-living environments is inherently chaotic. It is frequently plagued by missing values, anomalous readings, and asynchronous sampling intervals. A rigorous data preprocessing pipeline is therefore an absolute prerequisite before any meaningful feature extraction or machine learning can commence. The initial phase of preprocessing involves the identification and handling of missing data points, which commonly occur due to device

battery depletion, intentional removal of the wearable sensor, or background process terminations by the smartphone operating system [28]. For short-duration gaps, the system employs advanced interpolation techniques, carefully considering the temporal context of the missing segment. For instance, missing heart rate data over a brief five-minute interval can be reasonably approximated using spline interpolation, whereas replacing hours of missing physical activity data requires more sophisticated, behaviorally aware imputation models, such as using historical averages matched for the specific time of day and day of the week [29]. Following imputation, the system must synchronize the disparate data streams. Smartphone interaction logs are typically event-driven, generating data only when an action occurs, while wearable sensors produce continuous, fixed-frequency time series. To reconcile these differences, all data is resampled and aggregated into standardized temporal epochs. In this study, a one-hour epoch was selected as the optimal compromise, providing sufficient granularity to capture daily fluctuations while minimizing the computational overhead of overly dense data structures.

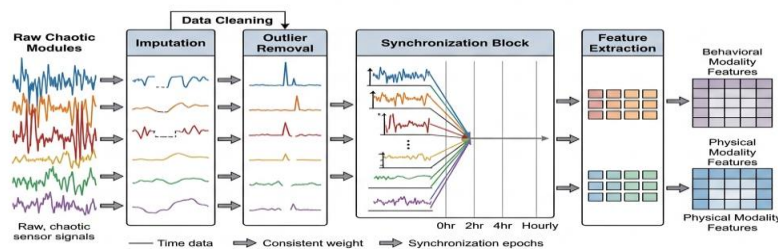


Figure 2: Multimodal Data Preprocessing Pipeline

4.2 Feature Engineering

Following data cleaning and synchronization, the subsequent critical step is feature engineering. The objective here is to transform the time-aligned epoch data into a compact set of highly descriptive statistical features that encapsulate the underlying behavioral and physiological patterns. Feature extraction is performed independently for both the behavioral and physical activity modalities over a rolling twenty-four-hour window, thereby capturing the complete diurnal cycle of the student. For the behavioral modality, features are primarily derived from the temporal distribution of smartphone usage. We calculate the total duration of interaction, the variance in usage session lengths, and the ratio of daytime to nighttime screen activity. Additionally, entropy-based features are computed to quantify the regularity of application usage patterns, with high entropy indicating erratic digital behavior often associated with emotional dysregulation [30]. For the physical activity modality, statistical features such as the mean, median, standard deviation, and skewness of heart rate and step counts are calculated for each epoch. We also extract frequency-domain features from the heart rate variability data, which are established biological markers of physiological stress. A comprehensive summary of the primary features extracted from both modalities is detailed in the table below.

Table 1: Summary of Extracted Multimodal Features

Modality	Feature Category	Description of Features Extracted	Specific
Behavioral	App Usage	Total duration, distribution, app launch frequency	category

Behavioral	Screen Activity	Unlock count, session length variance, night usage ratio
Behavioral	Routine Regularity	Entropy of daily usage, similarity to weekly baseline
Physical	Kinematics	Total step count, sedentary duration, peak activity intensity
Physical	Cardiac	Resting heart rate, heart rate variability, diurnal variance
Physical	Sleep	Total sleep duration, sleep efficiency, awakening frequency

5. Machine Learning Models for Mental Health Screening

5.1 Algorithmic Selection and Baselines

With a robust set of engineered features, the focus shifts to designing and training machine learning algorithms capable of accurately classifying the mental health status of the students. The primary task is formulated as a supervised classification problem, where the objective is to predict whether a student exhibits symptomatic levels of stress, anxiety, or depression based on their multimodal feature vectors over a specified observation window. To establish a comprehensive analytical framework, a variety of algorithms ranging from classical statistical learning models to advanced ensemble techniques are evaluated. Traditional algorithms such as Logistic Regression and Support Vector Machines are employed as foundational baselines. While these models offer high interpretability and are less prone to overfitting on smaller datasets, they often struggle to capture the complex, non-linear interactions inherent in multidimensional behavioral data. To address this, tree-based ensemble methods, specifically Random Forests and Gradient Boosting Machines, are heavily utilized. These algorithms excel at handling heterogeneous feature sets and implicitly perform feature selection, making them highly suitable for the diverse array of variables extracted in this study [31]. The ensemble models construct multitude decision trees during training and output the mode of the classes for classification, thereby providing robustness against the noise inevitably present in passive sensing data.

5.2 Multimodal Fusion Strategies

The crux of the modeling phase lies in the strategic fusion of the behavioral and physical activity modalities. The manner in which these data streams are integrated fundamentally dictates the predictive capability of the system. This research rigorously compares three distinct fusion architectures: early fusion, late fusion, and an intermediate fusion approach utilizing deep neural networks. In the early fusion paradigm, all features from Table 1 are concatenated into a single, high-dimensional vector prior to model training. This method allows the chosen algorithm to explore cross-modal interactions from the outset; however, it risks overwhelming the model with redundant or dominant features from a single modality. Conversely, the late fusion architecture involves training separate, modality-specific models. One classifier is trained exclusively on smartphone behavioral features, while another is trained solely on wearable physical features. The output probabilities from these independent models are subsequently aggregated using a meta-classifier or a weighted voting mechanism [32]. This approach prevents modality dominance and enhances generalization but may overlook subtle, concurrent patterns that span both data sources. To overcome the respective limitations of early and late fusion, a deep learning-based intermediate fusion architecture is proposed. This model employs separate fully connected neural network branches to process the behavioral and physical features independently, learning optimal high-level representations for each. These localized representations are then merged at a deeper layer within the network, allowing the subsequent dense layers to learn the complex interplay between the modalities before making the final classification.

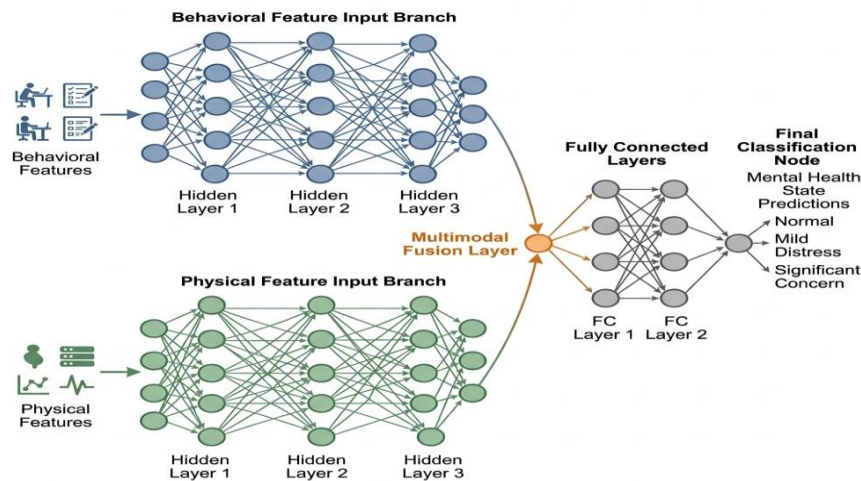


Figure 3 Intermediate Multimodal Fusion Architecture

6. Experimental Design and Evaluation Metrics

6.1 Dataset Description and Ground Truth Acquisition

To empirically validate the proposed multimodal screening framework, a comprehensive longitudinal study was conducted involving a cohort of undergraduate and graduate university students. Participants were recruited across various academic disciplines to ensure a representative sample of the student body. Upon enrollment, each participant provided informed consent and was equipped with the standardized wearable fitness tracker and the custom data acquisition smartphone application. The observation period spanned an entire academic semester, capturing periods of relative low stress at the beginning of the term as well as high-stress periods coinciding with midterm and final examinations [33]. The critical component of the experimental design is the establishment of robust ground truth labels for training and evaluating the machine learning models. To acquire these labels, participants were required to complete standardized clinical questionnaires bi-weekly through a secure web portal. The assessments included validated scales for stress, anxiety, and depression. Based on established clinical thresholds, the continuous scores from these questionnaires were binarized to create symptomatic and non-symptomatic classes for each psychological construct. This bi-weekly assessment interval was strategically chosen to balance the need for accurate temporal labeling with the necessity of minimizing participant survey fatigue.

6.2 Evaluation Methodology

The performance of the predictive models must be rigorously evaluated using metrics that accurately reflect their utility as clinical screening tools. Given that mental health datasets derived from student populations are frequently imbalanced, with the non-symptomatic class significantly outnumbering the symptomatic class, relying solely on standard classification accuracy is insufficient and potentially misleading. A model that universally predicts the majority class would achieve high accuracy while failing entirely to identify students requiring intervention. Consequently, the evaluation methodology heavily relies on Precision, Recall, and the F1-score. In the context of mental health screening, Recall is of paramount importance; it quantifies the system's ability to correctly identify all instances of psychological distress, thereby minimizing false negatives and ensuring that vulnerable students are not overlooked. Precision, conversely, measures the proportion of positive identifications that are actually correct, which is crucial for preventing the overburdening of university counseling services with false alarms. The F1-score provides a harmonic mean of Precision and Recall, offering a balanced single-metric assessment of model performance. Furthermore, Area Under the Receiver Operating Characteristic Curve is utilized to evaluate the models' discriminative capacity across various decision thresholds. To ensure the generalizability of the results and prevent data leakage, evaluation is conducted using strict leave-one-subject-out cross-validation, guaranteeing that the models are tested on entirely unseen individuals.

7. Results and Discussion

7.1 Performance Comparison of Modalities

The experimental results unequivocally demonstrate the superiority of the multimodal approach over unimodal baselines for the task of mental health screening. A comparative analysis was conducted across all three predictive targets: stress, anxiety, and depression. When relying solely on behavioral data extracted from smartphone usage, the models achieved moderate predictive capability, excelling primarily in detecting variations in social interaction and sleep routines. Models trained exclusively on physical activity data demonstrated strong performance in identifying somatic symptoms associated with depression, such as prolonged lethargy and disrupted heart rate variability, but struggled to differentiate between academic-induced stress and general physical exertion. By integrating both modalities using the proposed intermediate fusion deep learning architecture, the predictive performance experienced a substantial and statistically significant improvement across all metrics. The multimodal system was able to contextualize ambiguous signals, drastically reducing the false positive rate observed in the unimodal models. For example, a sudden drop in step count, which the physical-only model flagged as depressive behavior, was correctly contextualized by the multimodal model as intensive study behavior when coupled with a surge in productivity application usage on the smartphone. The quantitative results of this comparison are summarized in the table below, representing the cross-validated performance for predicting elevated stress levels.

Table 2: Performance Comparison of Unimodal and Multimodal Architectures

Modality Employed	Fusion Strategy	Precision	Recall	F1-Score
Behavioral Only	N/A	0.72	0.68	0.70
Physical Only	N/A	0.70	0.74	0.72
Multimodal	Early Fusion	0.76	0.75	0.75
Multimodal	Late Fusion	0.79	0.77	0.78
Multimodal	Intermediate Fusion	0.84	0.86	0.85

7.2 Clinical Implications and Limitations

The findings from this extensive evaluation carry profound implications for the future of psychiatric epidemiology and student support services. The ability to achieve an F1-score of 0.85 utilizing entirely passive, non-intrusive data collection methods suggests that continuous digital phenotyping can serve as a highly viable, scalable early warning system. By deploying such frameworks, university administration and campus health providers could transition from reactive treatment models to proactive intervention strategies. When the system detects a sustained negative trajectory in a student's multimodal feature profile, automated, personalized resources or invitations for counseling could be deployed long before a critical psychological crisis occurs. Furthermore, the granular nature of the extracted features allows clinicians to gain unprecedented insight into the specific behavioral and physiological disruptions accompanying a student's distress, facilitating more tailored and effective therapeutic interventions. Despite these highly promising results, several limitations within the current study must be acknowledged and addressed in future iterations. Primarily, the reliance on self-reported questionnaires for ground truth labels, while standard practice, inevitably introduces a degree of noise into the training data. The algorithms are fundamentally limited by the accuracy with which the students report their own psychological states. Additionally, while the intermediate fusion model demonstrates superior performance, its deep neural network architecture inherently functions as a black box, reducing the interpretability of the specific features driving the predictions. In clinical and educational settings, the ability to clearly explain why a screening system flagged an individual is often just as important as the accuracy of the flag itself. Finally, the algorithms were trained on a specific university demographic; variations in cultural norms, socioeconomic status, and institutional environments necessitate rigorous external validation before such systems can be universally deployed.

8. Conclusion and Future Directions

8.1 Summary of Findings

This comprehensive research paper investigated the feasibility and efficacy of a multimodal mental health screening framework utilizing passive behavioral and physical activity indicators collected from university students. By developing a robust, three-tiered system architecture, the study successfully harvested continuous, longitudinal data from standard consumer smartphones and wearable fitness trackers without burdening the user. The rigorous data preprocessing and feature extraction methodologies effectively translated raw, chaotic sensor signals into structured, clinically relevant matrices representing kinematic, cardiac, sleep, and digital interaction behaviors. Through extensive evaluation, the study conclusively demonstrated that fusing distinct modalities provides a far more accurate and nuanced representation of an individual's psychological state than relying on single data streams. Specifically, the implementation of a deep learning-based intermediate fusion strategy allowed the system to discover complex, cross-modal dependencies, culminating in high predictive accuracy for identifying symptomatic stress, anxiety, and depression. This approach successfully mitigated the high false positive rates characteristic of isolated digital phenotyping, proving its potential as a reliable, scalable early warning mechanism for educational institutions facing escalating mental health crises.

8.2 Future Research Trajectories

Moving forward, the field of multimodal mental health screening must prioritize the translation of these predictive models into actionable, real-world interventions. Future research should focus on the development of closed-loop systems, where the predictive insights generated by the algorithms are directly utilized to deliver personalized, just-in-time adaptive interventions directly to the student's device. Furthermore, enhancing the explainability of the complex fusion models is critical for fostering trust among clinicians and users. Integrating explainable artificial intelligence techniques to elucidate the exact combinations of behavioral and physiological features that trigger a mental health alert will be a vital step toward clinical acceptance. Finally, expanding the scope of data acquisition to include diverse demographic cohorts across various global institutions will be necessary to ensure the fundamental algorithms are equitable, unbiased, and universally effective in safeguarding student wellbeing on a global scale.

References

- Liu, Z. Q., Xiong, Y. J., Yu, X., Lv, T., & Xu, H. Y. (2026). Female reproductive factors, genetic susceptibility and incident osteoporosis risk. *Climacteric*, 1-8.
- Jing F, Ye Y, Zhou Y, et al. Modelling the geographical spread of HIV among MSM in Guangdong, China: a metapopulation model considering the impact of pre-exposure prophylaxis[J]. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, 2021, 380(2214): 20210126.
- Xu, Z., Lu, Y., Tao, X., Cheng, W., Tian, J., Tang, W., & Wu, D. Modelling the geographic spread of HIV among MSM in Guangdong, China: a metapopulation model considering the impact of pre-exposure prophylaxis *Philosophical Transactions A*.
- Wang, G., Zhou, Y., Deng, K., Cheah, W. P., Qu, R., Ren, J., & Shen, L. (2026). SurgCoT: Advancing Spatiotemporal Reasoning in Surgical Videos through a Chain-of-Thought Benchmark. *arXiv preprint arXiv:2604.20319*.
- Chen, Z., Zhu, Q., & Zeng, Y. (2026). MWED-Former: A Multi-Scale Weather-Aware Transformer for Emergency Department Demand Forecasting with Weather Bias Correction. *Fundamental Scientific Reports in Multidisciplinary Areas*, 2 (02), 132-145.
- Xi, Y., Zhao, Z., Zhou, Y., Yin, C., Li, Y., Xu, X.,... & Shen, W. (2026). Macrophage efferocytosis mediated by the TP63-RAC2 pathway promotes immunosuppressive remodeling in esophageal cancer. *Cell Reports Medicine*, 7 (1).
- Lei, C., Tao, X., & Xiao, Y. (2023, May). Dietary Iodine intake and related factors among secondary school students in Macao. In *Healthcare* (Vol. 11, No. 10, p. 1472). MDPI.
- GAO, T., LIANG, D., CHEN, X., & TAO, X. (2025). PREDICTIVE ANALYTICS IN PEDIATRIC MENTAL HEALTH: A MACHINE LEARNING APPROACH TO DEPRESSION DETECTION BY PHYSICAL ACTIVITY AND BEHAVIORAL HEALTH INDICATORS. *INTERCIENCIA*.

- Jiang, Z., Hou, Z., Yu, X., Zhao, Z., Bu, J., Li, C.,... & Wang, C. (2025). SPECT-derived myocardial perfusion and viability as predictors of response to left bundle branch pacing for cardiac resynchronization therapy. *Europace*, 27 (12), euaf301.
- Kelley, D. R., Snoek, J., & Rinn, J. L. (2016). Basset: Learning the regulatory code of the accessible genome with deep convolutional neural networks. *Genome Research*, 26(7), 990–999.
- Wang, G., Zhong, Z., Zhou, Y., Li, Y., Wu, E., Cheah, W. P., ... & Shen, L. (2026). X-PCR: A Benchmark for Cross-modality Progressive Clinical Reasoning in Ophthalmic Diagnosis.arXiv preprint arXiv:2604.20350.
- Yang, J., Chung, C. I., Koach, J., Liu, H., Navalkar, A., He, H.,... & Shu, X. (2024). MYC phase separation selectively modulates the transcriptome. *Nature Structural & Molecular Biology*, 31 (10), 1567-1579.
- Chen, Y., Sun, Y., Zhang, X., Chen, Y., Zhao, S., Shi, M., ... & Peng, M. (2025). A nomogram based on coagulation markers for predicting meige syndrome risk.*Frontiers in Aging Neuroscience*,17, 1665276.
- Li, S., Yu, X., Ma, X., Wang, Y., Guo, J., Wang, J.,... & Yin, S. (2024). Optimizing a classification model to evaluate individual susceptibility in noise-induced hearing loss: Cross-sectional study. *JMIR Public Health and Surveillance*, 10, e60373.
- Kabra, M., Shahi, P. K., Wang, Y., Sinha, D., Spillane, A., Newby, G. A., ... & Pattnaik, B. R. (2022). Nonviral base editing of KCNJ13 mutation preserves vision in an inherited retinal channelopathy.*bioRxiv*, 2022-07.
- Li, X., Lin, Y., He, W., Liu, R., Oliveira, A. L., Qian, T., ... & Hon, C. (2025). Enhancing the interpretation of spirometry: joint utilization of n-order adaptive fourier decomposition and deep learning techniques. *IEEE Transactions on Instrumentation and Measurement*.
- Zhang, Y., & Chen, Y. (2025, May). Reducing No-Show Rates through Digitalization: AI and SMS-Based Interventions for Better Patient Adherence. In *Proceedings of the 2025 International Conference on Digital Economy and Intelligent Computing* (pp. 123-127).
- Zhao, K., Wu, X., Han, G., Sun, L., Zheng, C., Hou, H.,... & Shi, Z. (2024). *Phyllostachys nigra* (Lodd. ex Lindl.) derived polysaccharide with enhanced glycolipid metabolism regulation and mice gut microbiome. *International Journal of Biological Macromolecules*, 257, 128588.
- Luo Y, Ma H, Sun J, et al. Inhibition of USAG-1 improved delayed graft function in renal transplantation[J]. *Transplant Immunology*, 2026: 102348.
- Zhao, K., Li, G., Li, C., Jiang, T., Wang, W., & Yang, Z. (2026). Identification Markers for *Salvia miltiorrhiza* and Its Close Relatives Based on Cell Wall Component Characteristics. *Engineered Science*, 40, 2122.
- Miotto, R., Li, L., Kidd, B. A., & Dudley, J. T. (2016). Deep patient: An unsupervised representation to predict the future of patients from the electronic health records. *Scientific Reports*, 6, 26094.
- Sinha, D., Tong, Y., Zepeda, M. A. F., Hanstad, G., Nelson, E. C., Theisen, C. O., ... & Gamm, D. M. (2023). Silica nanocapsules as a nonviral delivery platform for iPSC-RPE.*Investigative Ophthalmology & Visual Science*,64(8), 774-774.
- He, H., Sun, Z., Chen, X., Tao, X., Tao, M., Dong, D., ... & Qu, C. (2025). Exposure to volatile organic compounds and suicidal ideation: Insights from a US population-based study. *Journal of Affective Disorders*, 379, 194-203.
- Zhang, Y., & Chen, Y. (2025, July). The role of machine learning in reducing healthcare costs: The impact of medication adherence and preventive care on hospitalization expenses. In *Proceedings of the International Conference on Implementing Generative AI into Telecommunication and Digital Innovation 2025* (pp. 179-185).
- Wang, G., Li, Y., Chen, W., Ding, M., Cheah, W. P., Qu, R., ... & Shen, L. (2025, April). S³-mamba: Small-size-sensitive mamba for lesion segmentation. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 39, No. 7, pp. 7655-7664).

- Chai L, Li H, Zhao X, et al. Analysis of altered flowering related genes in a multi-silique rapeseed (*Brassica napus* L.) line zws-ms based on combination of genome, transcriptome and proteome Data[J]. *Plants*, 2023, 12(13): 2429.
- Ying, N., Wang, Y., Li, Z., Yu, X., Xu, X., Guo, Z., ... & Guo, T. (2026). Pathogenesis and role of ferroptosis mediated by Nrf2 signaling pathway in ischemic stroke. *Biochemical and Biophysical Research Communications*, 154363.
- Yin, L., Tong, Y., Xie, R., Zhang, Z., Islam, Z. H., Zhang, K.,... & Wang, B. (2024). Targeted NAD⁺ repletion via biomimetic nanoparticle enables simultaneous management of intimal hyperplasia and accelerated re-endothelialization: a proof-of-concept study toward next-generation of endothelium-protective, anti-restenotic therapy. *Journal of Controlled Release*, 376, 806-815.
- Xu, X., Bowtell, J., Fong, D. T., Young, W. R., & Williams, G. K. (2025). Kinematics of balance controls in people with chronic ankle instability during unilateral stance on a moving platform. *Scientific Reports*, 15(1), 1126.
- Yu, X., Zhang, Y., Yao, Y., Cho, W. C., & Shang, A. (2025). Biofilm-mediated immune dysregulation in chronic pulmonary diseases: mechanisms and clinical implications. *Frontiers in Microbiology*, 16, 1736384.
- Xu, X., Bowtell, J., Young, W. R., Fong, D. T., & Williams, G. K. (2026). Increased common corticospinal input during eyes-closed unilateral stance in people with chronic ankle instability. *Scientific Reports*.
- Chen, Z., Gao, P., Lee, Y., Barthelemy, J., Zhou, L., & Wang, L. (2025, June). Optimizing Efficiency and Visual-Textual Alignment for LLM-Based Radiology Report Generation. In *2025 IEEE International Conference on Multimedia and Expo (ICME)* (pp. 1-6). IEEE.
- Wang, G., Ren, J., Shen, L., Cheah, W. P., & Qu, R. (2026). M-MambaS: Multimodal Mamba for small lesion segmentation. *Pattern Recognition*, 113923.