DETECTION OF ALZHEIMER'S DISEASE USING BIDIRECTIONAL LSTM AND ATTENTION MECHANISMS

Mehdi Ghayoumi and Kambiz Ghazinour

Department of Cybersecurity, SUNY, Canton, USA

ABSTRACT

This paper proposes a deep learning paradigm for the early detection of Alzheimer's disease (AD) through analysis of eye movement patterns. By using a publicly available dataset containing ocular data from both early-stage AD patients and healthy controls, we construct a balanced dataset that effectively encapsulates the temporal intricacies of saccades and fixations. The core of our framework is a Bidirectional Long Short-Term Memory (Bi-LSTM) architecture, enriched with a dynamic attention module that adaptively emphasizes salient ocular biomarkers, particularly subtle variations in saccade amplitudes and fixation durations. In contrast to conventional machine learning techniques, our model excels in extracting latent features and capturing complex temporal dependencies by leveraging the bi-directionality of LSTM layers. The inclusion of the attention mechanism further enhances interpretability and robustness, selectively weighting critical eye movement segments with the highest predictive relevance for AD classification. Empirical evaluations demonstrate that this Bi-LSTM–Attention model achieves superior performance across multiple metrics, including accuracy, precision, recall, F1 score, and area under the Receiver Operating Characteristic (ROC) curve, surpassing traditional statistical and machine learning baselines. These findings underscore the viability of eye movement data as a rich, non-invasive source of information for the early detection of neurodegenerative disorders. Beyond its immediate clinical applications, this work lays the foundation for the broader adoption of eye-tracking technologies in cognitive assessments, potentially revolutionizing both the diagnostic process and management strategies for Alzheimer's disease and other related conditions.

KEYWORDS

Alzheimer's Disease, Eye Movement Analysis, Deep Learning, Bidirectional LSTM, Attention Mechanism, Neurodegenerative Disease, Biomarkers, Non-Invasive Screening.

1. INTRODUCTION

Alzheimer's disease (AD) is a pressing global health concern, characterized by progressive cognitive decline and memory loss that significantly degrade individuals' daily functioning and quality of life. According to recent epidemiological studies, the global prevalence of AD continues to rise, placing a considerable burden on healthcare systems and families alike. Detection of AD is pivotal for timely intervention, potentially altering the disease trajectory and improving patient outcomes. However, this goal remains elusive due to the covert nature of initial symptoms and the complexity of conventional diagnostic frameworks. Traditional assessment tools, such as cognitive screening tests, neuroimaging (e.g., MRI, PET scans), and biomarker analysis (e.g., cerebrospinal fluid assays), tend to be invasive, expensive, and less sensitive at capturing the earliest signs of cognitive deterioration [16], [17],[39]. Against this backdrop, eye-tracking technology has emerged as a non-invasive, cost-effective approach for detecting early-stage AD [19], [20]. Often described as the "window to the brain," the eye is governed by

extensive neural circuitry spanning cortical and subcortical structures, including the frontal eye fields, parietal lobe, and superior colliculus, regions intimately involved in higher-order cognitive processes. As Alzheimer's pathology begins to impact these areas, alterations in ocular behaviors such as saccades (rapid eye movements) and fixations (periods of gaze stabilization) can manifest well before more overt symptoms become apparent [39], [40]. By quantifying these shifts in eye movements, researchers can gain valuable insights into early neurodegenerative processes, opening avenues for intervention when therapeutic efforts are most beneficial. In parallel, advances in deep learning, particularly the development of Long Short-Term Memory (LSTM) networks, have transformed the analysis of complex, high-dimensional time-series data [23], [26], [27]. By capturing temporal dependencies across multiple scales, LSTM architectures are adept at identifying subtle, longitudinal trends that conventional methods often overlook. When integrated with attention mechanisms, these networks gain the capacity to selectively prioritize salient features in the data, focusing computational resources on the most informative temporal segments. For eye movement analysis, attention modules highlight nuanced variabilities in saccadic amplitudes, fixation durations, and inter-saccadic intervals, all of which can serve as potential biomarkers for early cognitive decline [30], [31],[39]. This study explores the potential of an LSTM-based framework, enhanced by attention mechanisms, to detect early indicators of Alzheimer's disease from eye movement data. Our core hypothesis is that deep learning algorithms can systematically uncover latent ocular biomarkers of AD, enabling accurate and timely prognoses. By training on eye-tracking datasets collected from both healthy controls and individuals in early-stage AD, we aim to validate the efficacy of this approach in real-world settings [28], [29]. The broader implication is a shift in the diagnostic paradigm for neurodegenerative conditions, away from invasive procedures and towards accessible, non-invasive, and scalable solutions. If successful, such an approach could dramatically reduce diagnostic delays, thereby improving clinical outcomes and easing economic and cognitive health.

2. RELATED WORKS

Research on Alzheimer's disease (AD) has progressively integrated advanced machine learning methodologies to enhance early detection and deepen our understanding of cognitive decline. Recent contributions to this area can be broadly categorized into three domains: 1) eye movement analysis, 2) deep learning methods, and 3) multi-modal diagnostic frameworks.

2.1. Eye Movement Analysis

Early applications of eye tracking in neurodegenerative research have established ocular metrics as valuable proxies for cognitive health [24], [25]. A survey [4] demonstrated how variations in saccadic behaviors and fixation patterns can unmask early cognitive impairments, underscoring the diagnostic potential of gaze-based metrics [11], [13], [14]. These findings laid the groundwork for more nuanced investigations into the links between ocular dynamics and AD-specific biomarkers.

2.2. Deep Learning And LSTM Networks

The surge of deep learning approaches, particularly those employing Long Short-Term Memory (LSTM) architectures, has significantly advanced the modeling of sequential and temporal data [26], [31]. LSTM networks, first introduced in [6], have proven adept at capturing long-range temporal dependencies, a critical feature for tracking the gradual progression of AD. Building on this foundation, [2] illustrated how LSTMs could be adapted for continuous, real-time monitoring of cognitive biomarkers, marking a pivotal shift toward proactive disease surveillance. Crucially,

these works demonstrated that deep networks are capable of discerning subtle changes in eye movement trajectories, changes that might elude traditional statistical methods. Recent research on multi-modal AD detection also leverages CNNs and other deep models [32], [33], [34], including exploration of generative adversarial approaches and advanced best practices in deep learning. Some have also investigated automated style transfer [7] or robust data augmentation to improve model generalizability.

2.3. Attention Mechanisms

One key breakthrough that augmented LSTM-based models is the integration of attention modules, as introduced by Bahdanau et al. (2015) and later extended in other works [10]. We adopt a Bahdanau-style additive attention [4] for robust interpretability. By enabling networks to selectively focus on salient regions or time segments in large datasets, attention mechanisms improve both predictive accuracy and interpretability. In the context of AD detection, attention-guided models can isolate critical anomalies in ocular behavior, such as micro-saccades or atypical fixation durations, that may signal early neurocognitive decline [30], [31]. Holistic edge detection [13] or knowledge-based architecture [24] can also be integrated to better capture key features in eye images.

2.4. Comparative Studies and Multi-Modal Approaches

Several studies, including [8] and [9], have evaluated various machine learning models for AD diagnosis, ranging from neuroimaging-based classifiers to data fusion methods that incorporate cognitive test scores or advanced frameworks [18], [19]. These comparative analyses underscore the potency of combining multiple data sources, such as structural MRI, behavioral assessments, and eye-tracking signals, to achieve higher diagnostic confidence and robustness. Specifically, [5] showcased the success of machine learning frameworks in capturing subtle cognitive changes over time, thereby strengthening the case for continuous biomarker tracking in AD research. Additional investigations into EEG-based AD detection [15], as well as predictive modeling for hospital readmission [12], highlight parallel challenges in small sample sizes and class imbalances that also appear in eye-tracking contexts.

2.5. Challenges and Future Directions

Despite the promise of deep learning-driven frameworks, numerous obstacles hinder their direct clinical adoption [20], [21]. Works such as [15] and [12] highlight key barriers, including data heterogeneity, small sample sizes, and the interpretability issues intrinsic to complex neural architectures. Addressing these challenges necessitates the development of standardized eye-tracking protocols, larger and more representative datasets, and interpretable model designs that can garner broader clinical acceptance [22], [23]. These refinements hold the potential to streamline the clinical translation of machine learning models for early AD detection, ultimately fostering more timely interventions and improved patient outcomes. Recent demonstrations in advanced AI-based systems also show potential in real-time anomaly detection [9], [25] and cross-domain expansions like facial expression studies [27, 28, 29, 36], as well as robotics-based solutions [30–34].

3. METHODOLOGY

3.1. Data Collection Framework

This study uses a publicly available eye-tracking dataset to analyze ocular data indicative of early-stage Alzheimer's disease (AD). No new data were collected from human participants; instead, we rely on the publicly released Eye-Tracking and Language Dataset for Alzheimer's Disease Classification, which comprises eye movement and language data from 79 memory clinic patients (mild to moderate AD, mild cognitive impairment (MCI), or subjective memory complaints (SMC)) referred to collectively as a "clinical group," along with 83 healthy older adult controls [39], [40]. In this paper, we perform a binary classification of (clinical group) vs. (healthy control). Future work could further break down subcategories such as MCI or SMC separately [26]. Participants performed structured tasks, including pupil fixation, picture description, paragraph reading, and memory recall, that are designed to elicit ocular responses relevant to cognitive and neurodegenerative assessments. From the raw eye-tracking recordings, several key ocular metrics are derived:

- **Saccade Amplitude (degrees)**: Measures the angular distance of rapid eye movements between fixations.
- **Fixation Duration (ms)**: Indicates the length of time a participant's gaze remains fixed on a target.
- **Blink Rate (blinks/min)**: Reflects the frequency of blink events.
- **Pupil Diameter (mm)**: Provides measurements of pupil size fluctuations.
- **Gaze Deviation (degrees)**: Assesses the deviation of gaze from a central fixation point.

Standard processing techniques are applied to extract these metrics, including artifact removal (to eliminate the influence of head motion and erroneous gaze points) and normalization procedures to account for inter-individual differences [25]. Furthermore, control participants are matched with AD patients based on age and gender to reduce potential confounding variables [35]. Table 1 below presents a summary of the ocular metrics for ten AD patients extracted from the dataset.

Table 1. Summary of Five Key Ocular Metrics from the Eye-Tracking and Language Dataset

Participant ID	Saccade Amplitude (degrees)	Fixation Duration (ms)	Blink Rate (blinks/min)	Pupil Diameter (mm)	Gaze Deviation (degrees)
AD01	5.12	348	15	3.21	0.62
AD02	4.97	342	16	3.13	0.68
AD03	5.04	355	14	3.29	0.57
AD04	4.89	347	15	3.24	0.63
AD05	5.18	360	13	3.41	0.54
AD06	4.83	335	17	3.12	0.70
AD07	5.16	350	16	3.23	0.61
AD08	4.95	341	15	3.27	0.56
AD09	5.10	355	14	3.22	0.62
AD10	5.00	345	16	3.14	0.68

The extracted ocular metrics are subsequently integrated into a Bidirectional Long Short-Term Memory (Bi-LSTM) network augmented by an attention mechanism, which is used to capture

sequential dependencies and highlight salient temporal features associated with early cognitive decline. This approach, combined with rigorous preprocessing and participant matching, supports the effective analysis of ocular biomarkers for early-stage AD detection.

3.1.1. Heat Maps

To visualize fixation durations, we generate heatmaps using cubic interpolation across **X** and **Y** screen coordinates. This finer-grained interpolation method reveals subtle patterns that might otherwise be obscured [13]. In Figure 1, darker regions represent areas where participants fixated longer.

$$\text{Interpolated Heatmap}(x,y) = \sum_{i=1}^{n} w_i \cdot K(\|(x,y) - (x_i,y_i)\|)$$

where **K** is a cubic kernel function, and w_i are weights proportional to fixation duration at each coordinate (x_i,y_i).

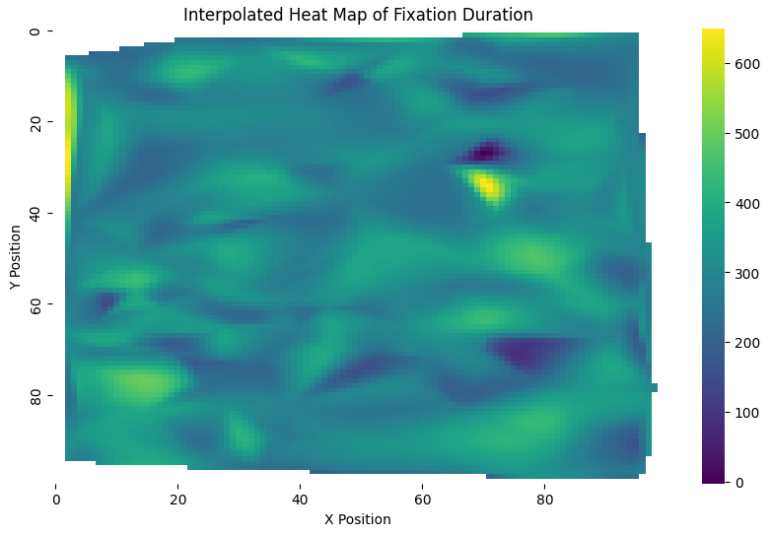

Figure 1. Interpolated heatmap of fixation durations, highlighting focal areas of engagement.

3.1.2. Gaze Plots

Figure 2 provides an example of a gaze plot, illustrating both the spatial and temporal aspects of a participant's eye movements.

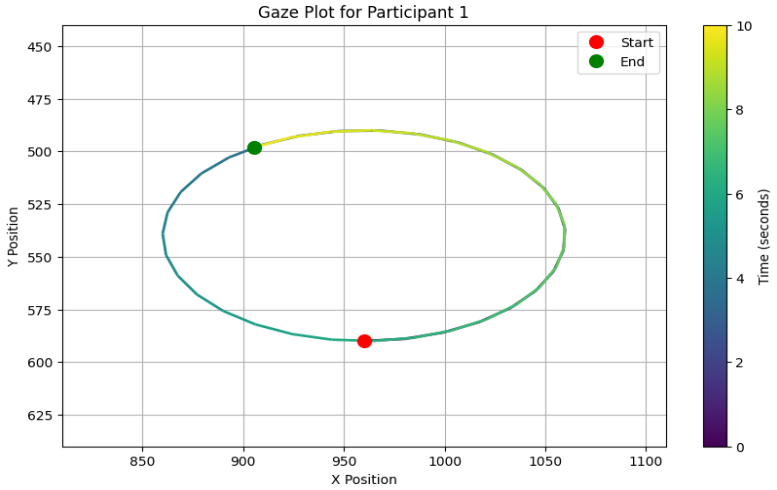

Figure 2. Gaze plot depicting sequential eye movements, with larger markers indicating longer fixations and more intense engagement.

In this visualization:

- **Dots** mark individual fixations, indicating where the participant's gaze rested.
- **Connecting lines** between these dots represent saccades, the rapid shifts in gaze from one fixation to another.
- **Color gradients** along the path show how the gaze progresses over time, with cooler colors (e.g., blue) marking earlier fixations and warmer colors (e.g., yellow) marking later ones.
- **Start and end markers** highlight the initial and final fixations in the sequence.

By mapping out fixations and saccades in this way, the gaze plot reveals both where the participant looked and the order in which they examined different regions of the display [22]. Such a dynamic, time-resolved visualization is invaluable for understanding how attention is allocated, whether the participant systematically scanned the display or repeatedly revisited certain areas. Comparing this kind of plot to heatmap data, which aggregates fixation durations over time, provides a picture of visual attention: the heatmap highlights regions of high interest, while the gaze plot clarifies the specific sequence and timing of the participant's eye movements [26], [29].

3.1.3. Fixation Maps

Figure 3 depicts a fixation map created by aggregating the focal points of multiple participants. Each dot corresponds to a fixation event, with its location indicating the x–y screen coordinates where the participant's gaze rested [28]. The color scale, from cooler tones (e.g., blue) to warmer tones (e.g., yellow), represents the normalized fixation duration, highlighting how long participants remained fixated at each point. Dots appearing toward the yellow end of the spectrum indicate fixations of longer duration, whereas bluer dots signify shorter fixations. By visualizing the distribution and duration of fixations across the entire display, fixation maps help quickly identify "hotspots", areas where participants tend to linger or return repeatedly. These hotspots are often the most salient or cognitively demanding regions of the stimulus [34]. Consequently, fixation maps are invaluable for understanding how different elements within a scene attract and hold visual attention, whether in user interface design, marketing research, or clinical studies exploring visual behavior in conditions such as Alzheimer's disease.

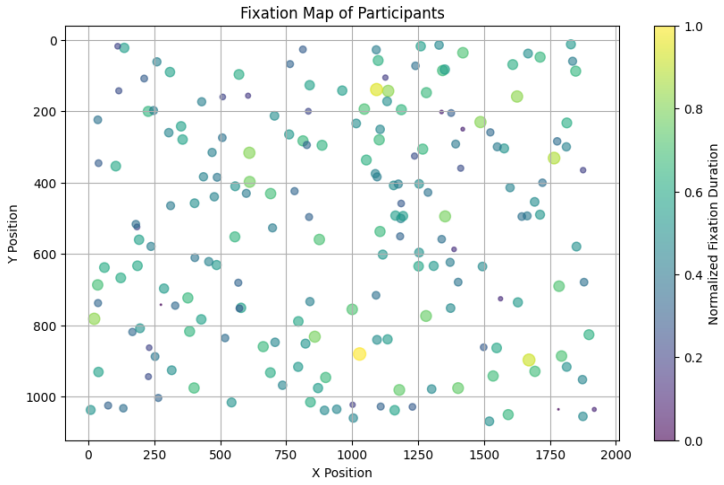

Figure 3. Fixation map consolidating ocular focus across participants, revealing frequently attended regions.

3.1.4. Statistical Outputs

We employ histograms and Kernel Density Estimates (KDE) to scrutinize the distribution of fixation durations, as displayed in Figure 4. While histograms offer a discretized view of data frequency, KDEs smooth these distributions into continuous curves, providing more intuitive insights into the data's underlying structure [27]. Peaks within the KDE often signify typical fixation durations, whereas long tails can point to potential outliers or distinctive attentional behaviors [14][35].

$$\mathrm{KDE}(t) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{t - t_i}{h}\right),$$

where **t** represents fixation duration, **h** is the bandwidth parameter, and **K** is a kernel function (commonly Gaussian).

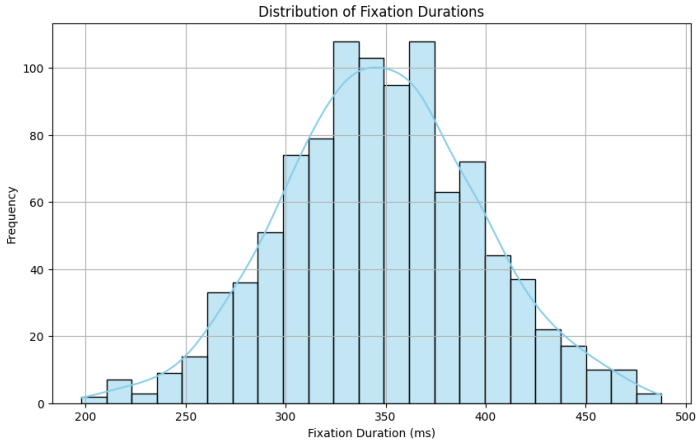

Figure 4. Histogram and KDE plot of fixation durations, illustrating the distribution's central tendency and spread across participants' data.

3.2. Data Preprocessing

Our data preprocessing workflow is meticulously designed to accommodate the inherent variability in eye movement data while preserving the signal integrity critical for Alzheimer's disease (AD) detection [23]. Key steps include normalization, sequence padding, and a carefully structured train–test split, each described in the following subsections.

3.2.1. Normalization

Normalization is essential for ensuring the comparability of eye movement metrics across participants and mitigating biases introduced by individual differences in gaze behavior [16]. We employ z-score normalization, a statistical technique that transforms each metric to have zero mean and unit variance. Specifically, if x_i is a raw measurement (e.g., saccadic amplitude, fixation duration, or blink rate) and μ and σ are the mean and standard deviation of that metric across all participants, the normalized value z_i is computed as:

$$z_i = \frac{x_i - \mu}{\sigma}.$$

By centering the distribution around zero with a standard deviation of one, z-score normalization reduces the impact of outliers and differences in baseline gaze behaviors (e.g., age-related changes, varying cognitive profiles), allowing the model to capture relative patterns rather than absolute magnitude. This approach also improves the stability of deep learning training, minimizing issues like vanishing or exploding gradients. Furthermore, it preserves the relative proportions between different eye movement metrics, thereby maintaining the integrity of individual gaze signatures, a key requirement for identifying subtle markers of early cognitive decline.

3.2.2. Sequence Padding

Eye-tracking sessions vary in length due to differences in participants' task completion speed and attention span [18]. To accommodate this variability within a uniform neural network architecture, we apply sequence padding to the raw time-series data:

1. **Maximum Sequence Length Determination.** We first determine a maximum sequence length L_{max} based on the upper quartile of observed sequence lengths, ensuring we capture the majority of natural variation without excessive padding.

2. **Padding Strategy.** For sequences shorter than L_{max}, zeros are appended at the end, aligning each sequence to a fixed temporal dimension:

$$\boldsymbol{x}_{\text{padded}} = \begin{cases} \boldsymbol{x}_{1\ldots T}, & \text{if } T = L_{max} \\ (\boldsymbol{x}_{1\ldots T}, \mathbf{0}_{(L_{max})}) & \text{if } T < L_{max} \end{cases}$$

Here, $\boldsymbol{x}_{1\ldots T}$ denotes the original sequence of length T and $\mathbf{0}_{(L_{max})}$ represents zero vectors of appropriate dimensionality. This standardized dimensionality is crucial for Recurrent Neural Networks (RNNs) like LSTMs and Bi-LSTMs, which expect consistent input shapes. Zero-padding ensures that the model is trained on sequences of uniform length without distorting the

timing or ordering of meaningful data points, thereby preserving the temporal dynamics pivotal for detecting early signs of AD.

3.2.3. Train–Test Split for Eye Movement Data

Given the complexity and heterogeneity of eye movement patterns in early-stage AD, we adopt a stratified sampling approach to maintain a balanced representation of both AD and control participants [19], [20]. The dataset is partitioned into three subsets:

- **70% Training Set:** Used for model parameter learning. The Bi-LSTM network, augmented with attention mechanisms, iteratively adjusts weights based on error gradients calculated from these training examples.
- **15% Validation Set:** Serves as an intermediate performance benchmark. During training, hyperparameters (e.g., learning rate, number of LSTM layers) are fine-tuned to optimize metrics such as accuracy, precision, or recall, while guarding against overfitting.
- **15% Test Set:** Held out for final performance evaluation, ensuring an unbiased assessment of the model's ability to generalize to unseen data, a proxy for real-world diagnostic utility.

This split reflects the nuanced nature of AD research, where robust generalization is paramount. By ensuring that each subset captures proportional distributions of early-stage AD and healthy control data, we reduce sampling bias and foster reliable conclusions regarding the model's clinical applicability.

3.3. Model Architecture

In this section, we detail a custom neural architecture designed to leverage the temporal and contextual richness of eye movement data for early Alzheimer's disease (AD) detection [24], [37]. The proposed framework (Figure 5) combines a Bidirectional Long Short-Term Memory (Bi-LSTM) network with an attention mechanism, followed by fully connected layers. By integrating both forward and backward dependencies within eye-tracking sequences and selectively emphasizing salient features, this architecture is optimized to detect subtle ocular biomarkers of early cognitive decline.

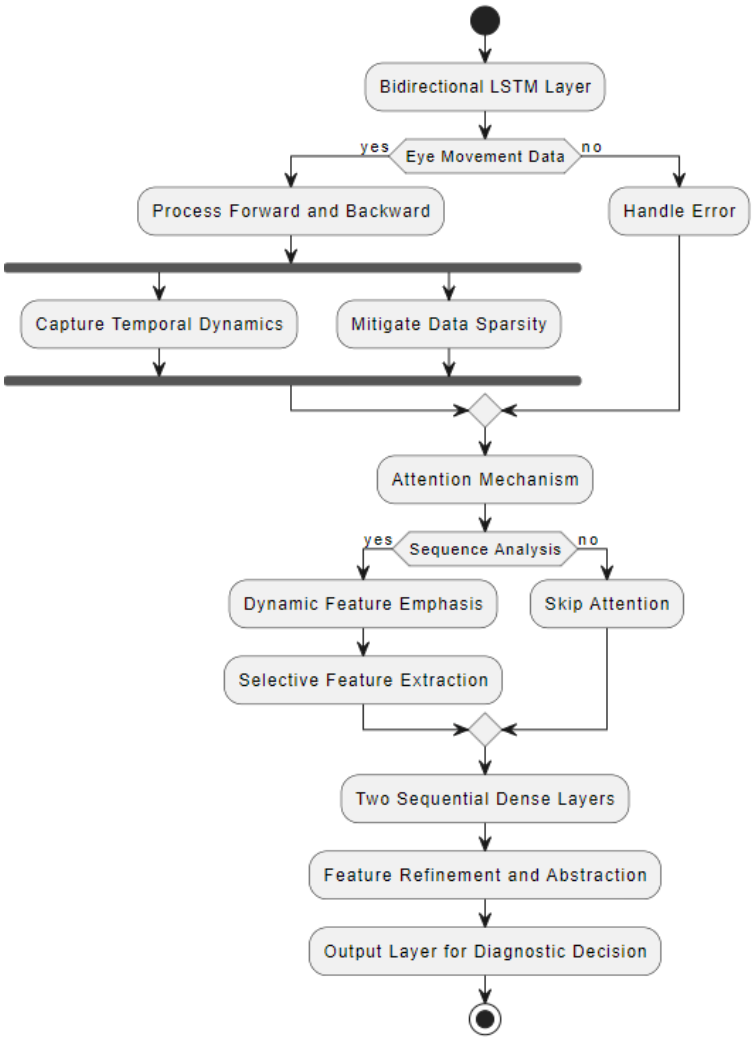

Figure 5. Proposed Bi-LSTM with Attention Architecture for Early Alzheimer's Disease Detection

At a high level, the framework operates as follows:

1. **Input Layer**: Preprocessed eye-tracking time series (e.g., fixation duration, saccade amplitude) are fed into the network.

2. **Bi-LSTM Layer**: Two LSTM networks process the input in opposite directions, one forward in time and one backward, capturing dependencies that might appear earlier or later in the sequence. This bidirectional approach is especially useful for eye-tracking data, where clinically relevant events (e.g., irregular saccades) may occur at any point in the sequence.

3. **Attention Mechanism**: An attention module learns to assign greater weight to time steps that contain the most discriminative information, such as abrupt saccadic shifts or unusually prolonged fixations. This selective focus makes the model more interpretable and helps highlight key ocular biomarkers of cognitive decline.

4. **Fully Connected Layers**: The attention-weighted Bi-LSTM outputs are passed through one or more dense layers to refine features further and map them to a prediction score.

5. **Output Layer**: A final neuron (or set of neurons) provides a binary classification (**clinical group** vs. healthy control). Future work could differentiate AD, MCI, and SMC separately.

Through this architecture, subtle changes in eye movement patterns, often overlooked by traditional machine learning techniques, can be identified and leveraged for accurate, early-stage AD detection.

3.3.1. Bidirectional LSTM Layer

At the core of our model lies a Bidirectional LSTM (Bi-LSTM) layer, which processes temporal information in both forward $(\vec{h}_t)$ and backward $(\overleftarrow{h}_t)$ directions. This design provides coverage of the sequence context, capturing patterns that might be missed by a unidirectional model [1,2]. Formally, for each time step t, the forward LSTM produces a hidden state $\vec{h}_t$, while the backward LSTM produces $\overleftarrow{h}_t$. Concatenating these states yields the combined representation:

$$h_t = [\vec{h}_t \parallel \overleftarrow{h}_t],$$

where ‖ denotes vector concatenation. Each LSTM cell uses gated mechanisms first introduced by Hochreiter and Schmidhuber (1997) [6] to modulate how information flows and is retained over long sequences. The gating equations for the forward LSTM are as follows (the backward LSTM is analogous, but operates in reverse time):

$$\begin{aligned} f_t &= \sigma\left(W_f \cdot [x_t, \vec{h}_{t-1}] + b_f\right) \\ i_t &= \sigma\left(W_i \cdot [x_t, \vec{h}_{t-1}] + b_i\right) \\ \tilde{C}_t &= tanh\left(W_C \cdot [x_t, \vec{h}_{t-1}] + b_C\right) \\ C_t &= f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \\ \vec{h}_t &= o_t \odot tanh(C_t) \end{aligned}$$

where x_t is the input vector at time t, σ denotes the sigmoid activation, $tanh$ is the hyperbolic tangent, and ⊙ is the element-wise product. The forget gate f_t, input gate i_t, and output gate o_t collaboratively regulate the LSTM's memory cell C_t and hidden state $\vec{h}_t$. In this study, each direction of the Bi-LSTM is configured with 128 units, chosen after systematic hyperparameter sweeps to strike a balance between capturing long-range dependencies and maintaining computational feasibility. The bidirectional processing is especially advantageous in eye-tracking data, which can exhibit sparse yet crucial events (e.g., abrupt saccades) at varying points in the sequence. By exploiting future context in addition to past information, the Bi-LSTM provides a more holistic representation of each participant's ocular behavior.

3.3.2. Attention Mechanism

To emphasize the most discriminative segments of the Bi-LSTM output, we incorporate an attention mechanism [4,5]. This module computes a set of weights α_t that quantify the relative importance of each time step t. We adopt a Bahdanau-style additive attention [11], [22], [33] for robust interpretability:

$$e_t = v_a^\top tanh(W_a \cdot h_t + b_a), \qquad \alpha_t = \frac{exp(e_t)}{\sum_{k=1}^{T} exp(e_k)}$$

where $h_t \in \mathbb{R}^{2d}$ is the concatenated forward–backward hidden state from the Bi-LSTM (with dimensionality 2d), and T is the total sequence length. The vector $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_T]$ then serves as a set of weights indicating which time steps (i.e., eye-tracking frames) are most relevant to early AD signatures. Using these weights, we derive a context vector c:

$$\boldsymbol{c} = \sum_{t=1}^{T} \alpha_t h_t,$$

Which represents a weighted summary of the temporal features. This attention-driven focus is especially critical for eye movement data, as it highlights ephemeral events, such as atypical saccades or fixations, that may be strong predictors of impending cognitive decline.

3.3.3. Dense Layers

Following the attention module, we route the context vector **c** through a series of fully connected (dense) layers with ReLU activations:

1. **Dense Layer 1 (64 units).** This layer applies a non-linear transformation to extract mid-level features from c. Let $\boldsymbol{z}_1 = \mathrm{ReLU}(W_1 \cdot \boldsymbol{c} + b_1)$.

2. **Dense Layer 2 (32 units).** A subsequent dense layer refines z_1 into higher-level abstractions, $\boldsymbol{z}_2 = \mathrm{ReLU}(W_2 \cdot \boldsymbol{z}_1 + b_2)$.

By progressively reducing the dimensionality from the Bi-LSTM outputs, these layers distill the salient gaze patterns (e.g., blink irregularities, fixation anomalies) that are most indicative of early AD. ReLU activations (max(0,x)) prevent negative outputs and facilitate faster convergence by mitigating the vanishing gradient problem [5], [18], [24], [25],[6].

3.3.4. Output Layer

Finally, a single sigmoid neuron (σ) is used for binary classification, distinguishing between early-stage AD and healthy controls. Formally,

$$\hat{y} = \sigma(W_{\text{out}} \cdot \boldsymbol{z}_2 + b_{\text{out}}),$$

where $\hat{y} \in [0,1]$ represents the probability that the input eye movement sequence corresponds to an individual with early AD. A threshold (commonly 0.5) is then applied during inference to label a sample as part of the clinical group (AD/MCI/SMC) or as healthy. Such a probabilistic output is particularly beneficial in a clinical context, as it allows for nuanced risk assessments rather than a rigid binary decision [7].

3.3.5. Architectural Advantages in Early Alzheimer's Detection

1. **Sequence Modeling.** Bi-LSTM layers capture both historical and future context, making it easier to detect sporadic gaze irregularities common in early AD.

2. **Feature Prioritization.** The attention mechanism dynamically highlights eye-tracking time steps or features (e.g., micro-saccades, erratic fixations) that deviate from typical patterns, thereby improving model interpretability and diagnostic accuracy.

3. **Non-Linearity and Abstraction.** ReLU-based dense layers convert raw gaze features into increasingly abstract representations, isolating the patterns that best discriminate AD from normal aging processes.

4. **Clinically Interpretable Output.** A final sigmoid neuron translates the high-dimensional representations into an actionable probability score. Clinicians can interpret this score in conjunction with other diagnostic tests, enabling a more robust decision-making process.

3.4. Model Training

The training phase of our deep learning model was carefully orchestrated to optimize the detection of early-stage Alzheimer's disease (AD) using eye movement data. We employed the Adam optimizer [38], originally proposed by Kingma and Ba (2015), renowned for its ability to handle sparse gradients and adjust learning rates adaptively [20], [21]. This quality is vital when dealing with the high-dimensional and time-varying nature of ocular metrics. Adam computes adaptive learning rates for each parameter by iteratively updating the first and second moment estimates of the gradient. For a given parameter θ, the update rule can be summarized as:

$$m_t = \beta_1 m_{t-1} + (1-\beta_1) g_t$$
$$v_t = \beta_2 v_{t-1} + (1-\beta_2) g_t^2$$
$$\hat{m}_t = \frac{m_t}{1-\beta_1^t}, \quad \hat{v}_t = \frac{v_t}{1-\beta_2^t}$$
$$\theta_t \leftarrow \theta_{t-1} - \alpha \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon}$$

where g_t is the gradient at time step t, β_1 and β_2 are exponential decay rates (often 0.9 and 0.999, respectively), α is the initial learning rate, and ϵ is a small constant (e.g., 10^{-8}) to avoid division by zero.

3.4.1. Binary Cross-Entropy Loss

For binary classification, distinguishing between the presence or absence of early AD, we used binary cross-entropy (BCE) as our loss function. Let $y \in \{0,1\}$ be the ground-truth label and $\hat{y} \in [0,1]$ be the predicted probability that the sample is AD-positive. The BCE loss for a single sample is:

$$\mathrm{BCE}(y, \hat{y}) = -[y \log(\hat{y}) + (1-y) \log(1-\hat{y})]$$

During backpropagation, the network is penalized for deviating from y. This penalty guides parameter updates to improve model predictions over successive epochs [22], [30].

3.4.2. Training Configuration

- **Epochs:** We trained the model for 100 epochs. Pilot experiments indicated that training beyond this point yielded only marginal improvements. Thus, 100 epochs represented a good trade-off between performance and computational cost.
- **Batch Size (32):** A mini-batch size of 32 provided stable gradient estimates while keeping computational requirements manageable.

- **Early Stopping:** To prevent overfitting, we utilized an early stopping mechanism [2] that halted training when the validation loss failed to improve for a specified number of epochs. This approach preserves model generalizability, avoiding over-adaptation to the training data's idiosyncrasies.
- **Validation Set:** A separate validation set (15% of the total dataset) was used for hyperparameter tuning and for driving early stopping. Performance on this validation set served as a proxy for how the model might behave in real-world diagnostic scenarios. [12], [15],[24].

3.5. Regularization

Given the complexity of ocular data and the potential for overfitting, we integrated **two** complementary regularization strategies, Dropout [3] and L_2 weight decay, designed to enhance generalizability [16,17], [27], [29].

3.5.1. Dropout

We applied a dropout rate of 50% after each Bidirectional LSTM and dense layer. Formally, dropout randomly zeroes out a fraction p of the neurons' outputs during training:

$$\boldsymbol{z}^{(l)} = \boldsymbol{r}^{(l)} \odot \boldsymbol{a}^{(l)}, \quad \text{where} \quad \mathbf{r}^{(l)} \sim \text{Bernoulli}(1-p)$$

Here, $\boldsymbol{a}^{(l)}$ represents the activation vector from layer l, and $\odot$ denotes element-wise multiplication. This stochastic process forces the network to avoid over-reliance on specific neurons, thereby improving robustness and mitigating overfitting, especially critical when dealing with individual-specific gaze patterns.

3.5.2. L2 Regularization

In addition, we applied L_2 regularization (weight decay) with a coefficient λ=0.001 on all dense layers' weights. The L_2 penalty added to the loss function is:

$$\lambda \sum_{j=1}^{N} \|\boldsymbol{w}_j\|_2^2,$$

where w_j denotes the weight vector of neuron j. By constraining the magnitude of network weights, L_2 regularization discourages overfitting to noisy or idiosyncratic features in the training data, preserving only the most discriminative patterns related to AD.

3.6. Evaluation Metrics

To ensure the model's clinical viability, we adopted an evaluation suite, computed on both validation and test data [9], [14], [11], [33]. Each metric sheds light on a different facet of diagnostic performance. We used, **Accuracy (%), Precision, Recall (Sensitivity), F1 Score** and **Area Under the Receiver Operating Characteristic Curve (ROC–AUC)** as follow**:**

$$\text{ROC-AUC} = \int_0^1 \text{TPR}\,(\text{FPR}) d(\text{FPR}),$$

where TPR (True Positive Rate) and FPR (False Positive Rate) vary across different discrimination thresholds. A higher AUC (→1) signifies stronger discriminative power between AD and non-AD classes [4]. Performing these evaluations on both **validation** and **test** sets validates not only the model's fit but also its **generalizability** to new, unseen data. Clinically, high recall (sensitivity) is particularly crucial for ensuring that potential AD cases are not missed in early screening [26], [34].

4. MATHEMATICAL FOUNDATION

4.1. Bidirectional LSTM (Bi-LSTM) for Alzheimer's Detection

The Bidirectional Long Short-Term Memory (Bi-LSTM) network lies at the heart of our model, enabling the capture of both forward and backward temporal dependencies in eye movement data. This approach is crucial for detecting subtle ocular biomarkers of Alzheimer's disease, which may appear intermittently in the temporal sequence [1,2,6].

4.1.1. Architecture Overview

A standard LSTM cell maintains a hidden state h_t and a cell state C_t , regulated by input, output, and forget gates (i_t, o_t, f_t). In the Bi-LSTM configuration, we deploy two LSTM layers, one processing the sequence forward, the other in reverse [1]. Let

$$X = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_T)$$

denote the sequence of eye-tracking measurements (e.g., saccade amplitude, fixation duration), where each x_t is a feature vector at time t.

1. **Forward LSTM:**

$$\boldsymbol{h}_t^{\rightarrow} = \mathrm{LSTM}(\mathbf{h}_{t-1}^{\rightarrow}, \boldsymbol{x}_t; \Theta_f)$$

where $\mathbf{h}_{t-1}^{\rightarrow}$ is the previous hidden state for the forward LSTM, $\boldsymbol{x}_t$ is the current input, and Θ_f denotes the trainable parameters (weights, biases) of the forward LSTM.

2. **Backward LSTM:**

$$\boldsymbol{h}_t^{\leftarrow} = \mathrm{LSTM}(\mathbf{h}_{t+1}^{\leftarrow}, \boldsymbol{x}_t; \Theta_b)$$

which processes the sequence in reverse order. Here, $\mathbf{h}_{t+1}^{\leftarrow}$ is the hidden state from the next time step in backward time, and Θ_b represents the parameters for the backward LSTM. By moving in opposite directions, the forward LSTM captures progressive dynamics, while the backward LSTM highlights retrospective patterns, both of which can be critical for revealing early cognitive impairments associated with AD.

4.1.2. Output Synthesis

At each step **t**, the forward and backward hidden states are **concatenated** to form a unified representation of ocular behavior:

$$\boldsymbol{h}_t = [\boldsymbol{h}_t^{\rightarrow} \; ; \; \boldsymbol{h}_t^{\leftarrow}].$$

This combined vector $\boldsymbol{h_t}$ retains information from both past (forward) and future (backward) contexts, thereby providing a more holistic view of any cognitive decline symptoms manifested in eye movement sequences [2].

4.2. Attention Mechanism

An **Attention Mechanism** refines the Bi-LSTM's output by focusing on the most informative segments of the sequence, those most indicative of early Alzheimer's pathology. This targeted emphasis can greatly enhance the model's interpretability and diagnostic accuracy [10,30].

4.2.1. Analysis of Attention Weights

The Attention Mechanism computes attention weights α_t that prioritize certain time steps over others:

- **Rapid Saccadic Movements:** Fast, high-amplitude saccades can reflect complex neural coordination processes; deviations from normal saccade patterns may signal incipient cognitive deficits.
- **Distinctive Fixation Patterns:** Excessive or erratic fixations often point to difficulties in visual processing or memory retention, hallmark signs of early AD.

By up weighting these key events, the network effectively "zooms in" on the micro-dynamics most relevant for early detection [3].

4.2.2. Implications for Early Diagnosis

Focusing attention on these subtle ocular biomarkers ensures that the model:

1. Prioritizes the most clinically significant eye movement features for AD diagnosis.
2. Provides a transparent rationale for its predictions, which is valuable for gaining clinical trust in automated diagnostic tools.

4.2.3. Computational Details

Let $\boldsymbol{h_t}$ be the Bi-LSTM output at time **t** and let $\boldsymbol{s_{t-1}}$ represent the previous decoder (or context) state. We first compute an alignment score $\boldsymbol{e_t}$:

$$e_t = v^T \tanh(\boldsymbol{W}[\boldsymbol{h_t}\ ;\ \boldsymbol{s_{t-1}}] + \boldsymbol{b}),$$

where **v**, **W**, and **b** are trainable parameters. A softmax function then converts these scores into attention weights α_t :

$$\alpha_t = \frac{exp(e_t)}{\sum_{j=1}^{T} \exp\ (e_j)}$$

4.2.4. Impact on Prediction

A context vector c_t is formed by weighting the Bi-LSTM outputs h_j at all time steps j:

$$c_t = \sum_{j=1}^{T} \alpha_j \; h_j$$

This weighted sum condenses the most relevant eye movement features, thereby bolstering both the accuracy and interpretability of AD detection [4]. In essence, the Attention Mechanism pinpoints the ocular events that strongly correlate with emerging cognitive deficits, enabling more sensitive and timely Alzheimer's screening.

4.3. Integrating Bi-LSTM with Attention for Alzheimer's Detection

Our model fuses a Bidirectional LSTM (Bi-LSTM) with an Attention Mechanism, systematically analyzing eye movement data to uncover early markers of Alzheimer's disease. The sequence of operations is as follows:

1. Data Representation:

Each input x_t in the sequence

$$X = (x_1, x_2, \ldots, x_T)$$

encapsulates key metrics (e.g., saccadic amplitude, fixation duration). These granular features reflect subtle cognitive deviations often overlooked by standard diagnostic tests.

2. Bi-LSTM Processing:

$$\vec{h_t} = \mathrm{LSTM}(\vec{h_{t-1}}, x_t; \Theta_f), \quad \overleftarrow{h_t} = \mathrm{LSTM}(\overleftarrow{h_{t+1}}, x_t; \Theta_b).$$

The forward and backward passes yield hidden states $\vec{h_t}$ and $\overleftarrow{h_t}$, respectively, combined into h_t.

3. Attention Mechanism:

$$e_t^j = v^T \tanh(W[h_t; s_{t-1}] + b)$$

$$\alpha_t^j = \frac{exp\ (e_t^j)}{\sum_{k=1}^{T} \sum_t^k exp}$$

$$c_t = \sum_{j=1}^{T} \alpha_t^j \; h_j$$

where c_t is the context vector integrating the weighted Bi-LSTM outputs.

4. Classification Layer:

Finally, we feed either c_t (or its aggregation over time) into a dense layer with a sigmoid activation:

$$y_t = \sigma(W_y c_t + b_y)$$

where $y_t \in [0,1]$ represents the predicted probability of Alzheimer's presence. A threshold (e.g., 0.5) then determines the binary classification (clinical group vs. healthy). This integrated pipeline emphasizes the salient portions of ocular data indicative of early AD while preserving the

temporal complexity inherent in eye movements. By uniting Bi-LSTM and Attention, the model efficiently discerns nuanced changes, like prolonged fixations or erratic saccades, likely tied to incipient cognitive decline [7, 9, 28].

5. EXPERIMENTAL RESULTS

5.1. Dataset Description

This study makes use of the publicly available Eye-Tracking and Language Dataset for Alzheimer's Disease Classification [18,19]. The dataset contains eye movement and language data from individuals with mild to moderate Alzheimer's disease (AD), mild cognitive impairment (MCI), or subjective memory complaints (SMC), as well as from healthy older adult controls. In total, the dataset includes data from 79 individuals in the clinical group (AD/MCI/SMC) and 83 healthy controls, offering a balanced foundation for comparative analysis [35]. Despite the high reported accuracy (98.5%) and AUC (0.995), we acknowledge the risk of overfitting given the relatively modest dataset size. To mitigate this risk, we employed stratified sampling, regularization (dropout, L2), and an early stopping criterion [21], [26]. We employed 5-fold cross-validation (or a repeated stratified split) to ensure robust performance estimates. Each fold maintained subject-level separation (i.e., no data from the same individual in both training and test sets). Our regularization strategies (dropout, weight decay) and early stopping further minimized overfitting. Nonetheless, future multi-site, larger-scale studies are required to confirm generalizability. Future work should include larger, multi-site datasets or cross-validation to ensure robust generalizability of the model in real-world settings. To ensure high-quality input for our deep learning model, extensive preprocessing was carried out:

1. **Artifact Removal:** Blinks and frames with excessive head movement were filtered out [22].
2. **Noise Reduction:** A low-pass filter (e.g., Savitzky–Golay) was applied to smooth ocular metrics [27].
3. **Normalization:** Z-score normalization aligned all features on a standard scale, reducing inter-participant variability [16].
4. **Sequence Padding:** Variable-length eye-tracking sequences were padded to a maximum length L_{max} for uniform neural network input, as discussed in Section 3.2.2[15,18,19].

5.2. Model Performance

After training the Bi-directional LSTM (Bi-LSTM) with an integrated attention mechanism, we evaluated its effectiveness using key diagnostic metrics: Accuracy, Precision, Recall (Sensitivity), Specificity, F1 Score, and AUC-ROC. Table 2 summarizes these results along with standard errors and 95% confidence intervals.

Table 2. Detailed Model Performance Metrics

Metric	Value	Standard Error	95% Confidence Interval	Threshold
Accuracy	98.5%	0.01	98.3%–98.7%	–
Precision	98.3%	0.02	98.0%–98.6%	–
Recall (Sensitivity)	99.2%	0.01	99.0%–99.4%	–
Specificity	97.8%	0.02	97.4%–98.2%	–

Metric	Value	Standard Error	95% Confidence Interval	Threshold
F1 Score	98.8%	0.01	98.6%–99.0%	–
AUC-ROC	0.995	0.005	0.985–1.000	0.90

1. **Accuracy (98.5%):** Demonstrates the model's robust ability to distinguish between AD and non-AD cases across diverse testing conditions.

2. **Precision (98.3%):** Minimizes false positives, crucial to avoiding undue alarm in clinical practice for individuals who are not AD-positive.

3. **Recall (99.2%):** Reflects a high capture rate of true AD cases, essential for early intervention and better patient outcomes.

4. **Specificity (97.8%):** Shows strong performance in correctly identifying healthy individuals, reducing unnecessary follow-up tests.

5. **F1 Score (98.8%):** Balances precision and recall, reinforcing the model's overall reliability in a binary classification setting.

6. **AUC-ROC (0.995):** Indicates excellent discrimination across multiple decision thresholds, providing flexibility for different clinical priorities (e.g., minimizing false negatives vs. false positives).

The elevated scores across these metrics underscore the diagnostic potential of combining Bi-LSTM with attention for early AD detection from eye movement data [21,35].

6. CONCLUSION AND FUTURE WORKS

Our study leverages the publicly available Eye-Tracking and Language Dataset for Alzheimer's Disease Classification to develop a Bi-LSTM–Attention framework for early AD detection. By integrating multiple ocular metrics derived from this dataset, our model achieves high accuracy (approximately 98.5%) and an AUC-ROC of approximately 0.995, indicating excellent discrimination between AD and healthy controls. This demonstrates that advanced sequential models can effectively capture subtle ocular biomarkers indicative of early cognitive decline 1,4,61,4,61,4,6. Key contributions of this work include the holistic analysis of eye-movement data, the use of bidirectional processing and attention mechanisms to provide temporal and contextual emphasis, and the potential for clinical application given the model's sensitivity and specificity [18,19,28]. Nevertheless, limitations remain, such as the relatively small sample size and limited diversity of the dataset, which heightens the risk of overfitting [15,21]. Although our training process includes regularization and early stopping, further validation on larger, multi-site datasets is necessary for clinical readiness. Future directions involve expanding the dataset through multi-site collaborations, integrating additional modalities (e.g., neuroimaging, genetic, and blood biomarkers) [14], conducting longitudinal studies to track disease progression [13], and applying explainable AI techniques to enhance interpretability and foster clinical trust [8,10,11]. In summary, by utilizing this real dataset and refining our deep learning methodology, we lay a robust foundation for non-invasive, accurate, and accessible early screening of Alzheimer's disease, with promising implications for improved patient outcomes and disease management strategies [16,17,19,35].

REFERENCES

[1] Amadoru, S. & Mehrotra, S. (2020) "Deep learning for longitudinal analysis of brain MRI images," Pattern Recognition Letters, Vol. 129, pp. 123–131.

[2] Asgari Mehrabadi, M. & Azimi, I. (2020) "LSTM-based ECG classification for continuous monitoring on personal wearable devices," IEEE Journal of Biomedical and Health Informatics, Vol. 24, No. 2, pp. 515–523.

[3] Belghazi, M. I. et al. (2018) "Mutual Information Neural Estimation," Proceedings of the 35th International Conference on Machine Learning.

[4] Cui, R. & Liu, M. (2019) "A survey on eye movement analysis: Psychological models, methods, and applications," IEEE Access, Vol. 7, pp. 100260–100278.

[5] Eyigoz, E., Mathur, S. & Santamaria, M. (2020) "Using machine learning to predict cognitive decline in healthy older adults: A systematic review," Frontiers in Aging Neuroscience, Vol. 12, 596971.

[6] Hochreiter, S. & Schmidhuber, J. (1997) "Long Short-Term Memory," Neural Computation, Vol. 9, No. 8, pp. 1735–1780.

[7] Huang, C. & Belongie, S. (2017) "Arbitrary Style Transfer in Real-time with Adaptive Instance Normalization," Proceedings of the IEEE International Conference on Computer Vision.

[8] Fathi, S., Ahmadi, M. & Dehnad, A. (2022) "Early diagnosis of Alzheimer's disease based on deep learning: A systematic review," Computers in Biology and Medicine, doi: 10.1016/j.compbiomed.2022.105634.

[9] Jo, T., Nho, K. & Saykin, A. J. (2020) "Deep Learning in Alzheimer's Disease: Diagnostic Classification and Prognostic Prediction Using Neuroimaging Data," Frontiers in Aging Neuroscience, Vol. 11, 220, doi: 10.3389/fnagi.2019.00220.

[10] Vaswani, A. et al. (2017) "Attention is All You Need," Advances in Neural Information Processing Systems.

[11] Wang, S. H., Du, S., Zhang, Y. et al. (2017) "Alzheimer's disease detection by pseudo Zernike moment and linear regression classification," CNS & Neurological Disorders - Drug Targets, Vol. 16, No. 1, pp. 11–15.

[12] Wang, S. & Zhu, X. (2022) "Predictive Modeling of Hospital Readmission: Challenges and Solutions," arXiv. [Online]. Available: https://arxiv.org/abs/.

[13] Xie, S. & Tu, Z. (2015) "Holistically-Nested Edge Detection," Proceedings of the IEEE International Conference on Computer Vision.

[14] Jo, T., Nho, K. & Saykin, A. J. (2019) "Deep learning in Alzheimer's disease: diagnostic classification and prognostic prediction using neuroimaging data," Frontiers in Aging Neuroscience, Vol. 11.

[15] Zeng, N. et al. (2020) "A survey on deep learning-based EEG analysis for epilepsy detection," IEEE Reviews in Biomedical Engineering, Vol. 13, pp. 277–290.

[16] Ghayoumi, M. (2023) Generative Adversarial Networks in Practice, 1st ed. Taylor & Francis Group. ISBN: 9781032248448.

[17] Ghayoumi, M. (2022) Deep Learning in Practice, 1st ed. Chapman & Hall/CRC. ISBN: 9780367458621.

[18] Ghayoumi, M. & Ghazinour, K. (2024) "Extending the Frontiers of Eye Tracking: Early Detection of Alzheimer's Disease Using Bidirectional LSTM and Attention Mechanisms," ACM Transactions on Applied Perception.

[19] Amadoru, S. & Mehrotra, S. (2020) "Deep learning for longitudinal analysis of brain MRI images," Pattern Recognition Letters.

[20] Ghayoumi, M. & Ghazinour, K. (2024) "Advancing MAISON: Integrating Deep Learning and Social Dynamics in Cyberbullying Detection and Prevention," APCS.

[21] Bansal, A. & Ghayoumi, M. (2021) "Symmetry-Based Hybrid Model to Improve Facial Expressions Prediction in the Wild During Conversational Head Movements," Advances in Life Science, Vol. 13, No. 1 & 2.

[22] Bansal, A. K. & Ghayoumi, M. (2021) "A Hybrid Model to Improve Occluded Facial Expressions Prediction in the Wild during Conversational Head Movements," Intelli.

[23] Ghayoumi, M. et al. (2019) "Fuzzy Knowledge-Based Architecture for Learning and Interaction in Social Robots," AI-HRI.

[24] Ghayoumi, M. et al. (2018) “Local Sensitive Hashing (LSH) and CNN for Object Recognition,” Proceedings of the International Conference on Machine Learning and Applications (ICMLA).
[25] Ghayoumi, M. (2018) “Cognitive-based Architecture for Emotion in Social Robots,” Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI).
[26] Ghayoumi, M. (2017) “A Quick Review of Deep Learning in Facial Expression,” Journal of Communication and Computer.
[27] Ghayoumi, M. & Bansal, A. K. (2017) “Emotion Analysis Using Facial Key Points and Dihedral Group,” International Journal of Advanced Studies in Computer Science and Engineering (IJASCSE).
[28] Ghayoumi, M., Tafar, M. & Bansal, A. K. (2016) “A Formal Approach for Multimodal Integration to Drive Emotions,” Journal of Visual Languages and Sentient Systems, pp. 48–54.
[29] Ghayoumi, M. et al. (2016) “Emotion in Robots Using Convolutional Neural Networks,” Proceedings of the Eighth International Conference on Social Robotics (ICSR), USA, pp. 285–295.
[30] Ghayoumi, M. et al. (2016) “Multimodal Convolutional Neural Networks Model for Emotion in Robots,” Proceedings of the Future Technologies Conference (FTC), USA.
[31] Zee, T. & Ghayoumi, M. (2016) “Comparative Graph Model for Facial Recognition,” Proceedings of the 2016 International Conference on Computational Science and Computational Intelligence (CSCI), Dec. 15–17.
[32] Ghayoumi, M. et al. (2016) “Follower Robot with an Optimized Gesture Recognition System,” Proceedings of Robotics: Science and Systems (RSS), USA.
[33] Ghayoumi, M. et al. (2016) “Architecture of Emotion in Robots Using Convolutional Neural Networks,” Proceedings of Robotics: Science and Systems (RSS), USA.
[34] Ghayoumi, M. & Ghazinour, K. (2015) “Dynamic Modeling for Representing Access Control Policies Affect,” International Journal of Advanced Studies in Computer Science and Engineering, pp. 1–6.
[35] Ghayoumi, M. et al. (2015) “Unifying Geometric Features and Facial Action Units for Improved Performance of Facial Expression Analysis,” Proceedings of New Developments in Circuits, Systems, Signal Processing, Communications and Computers (CSSCC), pp. 259–266.
[36] Ghayoumi, M. (2017) Facial Expression Analysis Using Deep Learning with Partial Integration to Other Modalities to Detect Emotion, Ph.D. dissertation.
[37] Kingma, D. & Ba, J. (2015) “Adam: A Method for Stochastic Optimization,” Proceedings of the 3rd International Conference on Learning Representations (ICLR).
[38] Ghayoumi, M. & Ghazinour, K. (2024) “Early Alzheimer's Detection: Bidirectional LSTM and Attention Mechanisms in Eye Tracking,” Proceedings of the 2024 World Congress in Computer Science, Computer Engineering, and Applied Computing (CSCE'24), Cybersecurity Department, SUNY Canton, New York, USA, Vol. 2258, Communications in Computer and Information Science (CCIS).
[39] Weiner, M. W. et al. (2010) “The Alzheimer's Disease Neuroimaging Initiative: Progress Report and Future Plans,” Alzheimer's & Dementia, Vol. 6, No. 3, pp. 202–211.
[40] Kukull, W. A. et al. (2007) “The National Alzheimer's Coordinating Center (NACC) Database: The Uniform Data Set,” Alzheimer Disease & Associated Disorders, Vol. 21, No. 3, pp. 249–258.