TU Delft

Delft University of Technology

Suicide ideation detection based on documents dimensionality expansion

Esmi, Nima; Shahbahrami, Asadollah; Gaydadjiev, Georgi; de Jonge, Peter

DOI
10.1016/j.compbiomed.2025.110266

Publication date
2025

Document Version
Final published version

Published in
Computers in Biology and Medicine

Citation (APA)
Esmi, N., Shahbahrami, A., Gaydadjiev, G., & de Jonge, P. (2025). Suicide ideation detection based on documents dimensionality expansion. *Computers in Biology and Medicine*, *192*, Article 110266. https://doi.org/10.1016/j.compbiomed.2025.110266

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Computers in Biology and Medicine 192 (2025) 110266

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Suicide ideation detection based on documents dimensionality expansion

Nima Esmi [a,b,*], Asadollah Shahbahrami [c,b], Georgi Gaydadjiev [d], Peter de Jonge [e]

[a] Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence, University of Groningen, Groningen, The Netherlands
[b] Intelligent Systems Research Center, Khazar University, Baku, Azerbaijan
[c] Department of Computer Engineering, University of Guilan, Rasht, Iran
[d] Computer Engineering Laboratory, Delft University of Technology, Delft, The Netherlands
[e] Faculty of Behavioural and Social Sciences, University of Groningen, Groningen, The Netherlands

ARTICLE INFO

Keywords:
Informal document classification
Dimensionality expansion
Convolutional neural networks
Social media
Suicide ideation detection

ABSTRACT

Accurate and secure classifying informal documents related to mental disorders is challenging due to factors such as informal language, noisy data, cultural differences, personal information and mixed emotions. Conventional deep learning models often struggle to capture patterns in informal text, as they miss long-range dependencies, explain words and phrases literally, and have difficulty processing non-standard inputs like emojis. To address these limitations, we expand data dimensionality, transforming and fusing textual data and signs from a 1D to a 2D space. This enables the use of pre-trained 2D CNN models, such as AlexNet, Restnet-50, and VGG-16 removing the need to design and train new models from scratch. We apply this approach to a dataset of social media posts to classify informal documents as either related to suicide or non-suicide content. Our results demonstrate high classification accuracy, exceeding 99%. In addition, our 2D visual data representation conceals individual private information and helps explainability.

1. Introduction

Text analysis is a common Natural Language Processing (NLP) task that involves analyzing one-dimensional (1D) textual data based on its content. This analysis can occur at various levels, ranging from individual words and sentences to paragraphs and entire documents. Document-level analysis has applications across diverse domains [1,2], including customer feedback analysis, sentiment analysis [3,4], spam detection [5–7], the identification of cyberbullying and toxic comments [8], detection of fake accounts and fake news [9–12], and the recognition of mental disorders, such as depression and suicidal ideation [13–15]. Document-level analysis presents unique complexities. Beyond local features, such as individual words, contiguous word sequences, and word or phrase positions within a document, this type of analysis must also capture global concepts, including the document's document-related topic or the author's general perspective on a given subject. Documents analyzed in these contexts frequently originate from informal communication platforms, often characterized by colloquial language, abbreviations, slang, and extensive use of emojis. Such documents may also include elements of sarcasm, irony, or culturally specific references, all of which require precise explanation to accurately convey intended meanings. Moreover, individuals' intentions, such as suicidal thoughts, may not be overtly expressed in social media texts; instead, they may subtly communicate distress or emotional confusion [16]. Explaining these implicit cues presents a significant challenge. There is a crucial need to demonstrate the capacity to detect subtle indicators, including shifts in language patterns, fluctuations in sentiment, or changes in engagement behaviors, to effectively identify individuals' tendencies [17–19]. Additionally, explaining and understanding the diagnostic outputs produced by AI models can be challenging. In sensitive areas, such as diagnosing mental disorders, it is vital to balance respecting individuals' privacy with the need for timely intervention and support. Ethical considerations in model development and deployment require a cautious approach, ensuring that privacy is maintained while striving to minimize harm and extend support to those in need [20,21].

Common deep learning methods for document-level analysis involve several steps, including pre-processing, word embedding, model training, and testing [22,23]. During pre-processing, tasks such as removing stop words (e.g., "the", "and", "is"), cleaning special characters, and disregarding other signs are performed. Popular word embedding techniques, such as Word2Vec, GloVe, and FastText, convert words into dense vector representations (embeddings) that capture semantic meanings and relationships among words [24]. In the model training phase, an appropriate deep learning architecture is selected for

* Corresponding author at: Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence, University of Groningen, Groningen, The Netherlands.
E-mail address: n.esmi.rudbardeh@rug.nl (N. Esmi).

https://doi.org/10.1016/j.compbiomed.2025.110266
Received 20 November 2024; Received in revised form 5 April 2025; Accepted 22 April 2025
Available online 13 May 2025
0010-4825/© 2025 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

text classification. For 1D data classification, architectures like Recurrent Neural Networks (RNNs) including Bidirectional Long Short-Term Memory (BiLSTM) and 1D Convolutional Neural Networks (1D CNNs) are commonly used [25]. However, these models face limitations at the document level, as they struggle to capture sequential and long-range dependencies between words due to issues like vanishing and exploding gradients, limited short-term memory in RNNs, the sequential processing nature of RNNs, and the fixed context window of 1D CNNs. Such challenges often lead to reduced classification accuracy. To address these limitations, hybrid and transformer-based models have been developed, though these introduce greater implementation complexity [26].

Moreover, in applications that involve user opinions and emotions – such as analyzing customer feedback or diagnosing mental disorders – implicit structures like metaphors and allusions add complexity to document-level analysis. Notably, online content frequently includes more than just text; people use signs like emojis and emoticons to express feelings and opinions. While these signs can provide valuable sentiment insights, they are often removed during pre-processing. Dimensionality transformation offers a promising approach to overcome these challenges.

Dimensionality transformation involves converting data from one domain to another. Dimensionality reduction simplifies models and reduces noise, whereas dimensionality expansion enables the capture of more complex relationships in the data, making it easier for machine learning algorithms to detect patterns [27]. This technique has diverse applications, including speech recognition [28], financial trading analysis [29], electroencephalography analysis [30], and DNA sequence detection [31].

In this paper, we employ dimensionality expansion for informal document classification. In our proposed model, signs such as emojis and emoticons are first separated from the text. Then, specific data cleaning processes are applied to the textual content. Both signs and words are transformed into numerical values and subsequently into visual representations. Finally, we fuse the representative images of emojis and words for each document into a single image, performing multimodal data fusion that integrates signs with textual data.

For classification, we utilize existing 2D pre-trained CNN models. The key contributions of our work are summarized as follows:

- By applying dimensionality expansion to 1D data and converting them into images, we utilize existing 2D pre-trained CNN models such as AlexNet [32], ResNet-50 [33], and VGG-16 [34], and with minimal time investment in fine-tuning operations. This approach eliminates the need to design and train new models from scratch, thereby reducing associated learning costs. Utilizing 2D CNN models enables simultaneous feature extraction at both local and global levels. At the local level, the models capture edges, textures, and corners, which represent tokenized words. At the global level, they capture the overall shape of objects, reflecting the author's overarching perspective. Additionally, the influence of signs on classification is incorporated. As a result, compared to other common models, the accuracy in diagnosing mental disorders has increased significantly, reaching more than 99%.
- By employing dimensionality expansion, we provide a visual representation of fused text and symbols. This approach preserves personal private data. In addition, it can be used as an auxiliary tool for result explanation.

The remainder of this article is organized as follows. Section 2 presents a comprehensive literature review, followed by essential information and fundamental concepts about two-dimensional pre-trained CNNs in Section 3. The proposed model is introduced in Section 4, and experimFental results are discussed in Section 5. Finally, we conclude the article in Section 6.

2. Related work

Various virtual environments, such as social networks, provide valuable opportunities for extracting and analyzing shared content [35], enabling an in-depth exploration of feelings, emotions, and tendencies such as aggression [36], cyberbullying [37–39], depression [17,40,41], and suicide [42–45]. Given that suicide attempts may be life-threatening, the literature on suicide risk assessment and diagnosis through the analysis of content produced by individuals on social networks using deep learning methods is discussed below. From this perspective, the literature can be divided into the following categories of models that analyze one-dimensional social network text data: Recurrent Neural Networks (RNNs) [46–48], Convolutional Neural Networks (CNNs) [49–52], transformers, and hybrid-based solutions [53–57]. Table 1 presents a selection of these models from the past seven years, organized by accuracy.

A Long Short-Term Memory (LSTM) network is a type of RNN that can learn long-term dependencies between data sequences, making it suitable for extracting the information and sentiments embedded in texts. In [47], feedback connections were incorporated in addition to the standard LSTM feed-forward structure, allowing previously collected information to be stored while redundant data is removed. In this model, attention was given to sentences at the word level, and a bag-of-words approach was used to extract statistical features at the word level for suicide ideation diagnosis. In [46], 16 different emotion-related factors were considered to prevent suicide, where an increase in the intensity of negative emotions was indicative of a higher suicide risk. Features were extracted using unigram and bigram techniques, word embeddings, and term frequency–inverse document frequency (TF–IDF). Sawhney et al. [58] introduced a transformation stage before their LSTM model to learn the semantic characteristics of social media content. Similarly, Zhang et al. [59] utilized a transformer model, with the added advantage of capturing long-term dependencies and contextual information.

Gaur et al. improved the prediction accuracy of suicide-related cases by combining time-variant approaches with CNNs [54]. Clinical data, such as mental health records and medications, were used as supplementary information [60]. In [55], laughter and sadness emoticons were used as weighting identifiers for non-suicidal and suicidal cases, respectively. Tadesse et al. [56] applied word embedding techniques before the learning stage, mapping words to a vector space of real numbers. A CNN model was then used to address the limitations of feed-forward networks in pattern retention. Naseem et al. [61] presented a model for detecting suicidal tendencies on social networks by combining transformers, utilizing a Longformer within the transformer structure to capture word features. Additionally, a text tensor was applied to extract text features and syntactic, semantic, and sequential context information. In [62], knowledge graphs were used to identify concepts related to suicide domains. Cao et al. [63] constructed a knowledge graph combined with deep neural networks to create implicit graph connections between text and images, incorporating personal information such as gender, age, and geographical location, which posed privacy concerns for users' data. Recent transformer-based models employ independent architectures that are generally more complex than CNN and RNN models, achieving relatively higher accuracy. For example, [64] used Generative Pre-trained Transformer 3 (GPT-3) and the Robustly optimized BERT pretraining approach (RoBERTa) for text analysis. In [65], a model based on an independent transformer architecture that integrates sentiment features with text was introduced.

3. Two dimensional pre-trained CNNs

There are several types of 2D pre-trained CNN models, including AlexNet [32], ResNet-50 [33], and VGG-16 [34]. These models consist

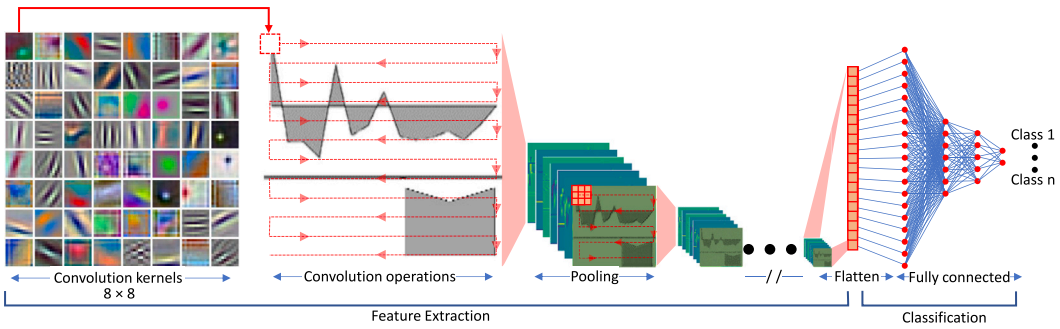

Fig. 1. A general overview of the classification operations performed by 2D pre-trained CNN models; includes the use of pre-trained convolution kernels for feature extraction, pooling to reduce dimensionality, and the utilization of flattened and fully connected layers to carry out the classification process.

Table 1
Comparison between some recent researches that have used deep learning models to detect suicide ideation in social networks, which are sorted based on accuracy.

Ref.	Year	Model	Dataset	Accu.	F1
[65]	2024	Transformer	Reddit	99.34	—.—
[64]	2023	Transformer	Twitter	97.40	—.—
[50]	2019	CNN	Reddit	95.40	96.60
[53]	2022	CNN-LSTM	Reddit	95.00	95.00
[56]	2020	LSTM-CNN	Reddit	93.80	93.40
[62]	2021	BERT-LSTM	Sina Weibo	91.25	91.23
[57]	2022	LSTM-CNN	Reddit	90.30	92.60
[51]	2021	CNN	Twitter	87.00	91.50
[46]	2022	LSTM	Facebook	80.00	—.—
[54]	2021	LSTM-CNN	Reddit	78.00	79.00
[52]	2018	CNN	Twitter	74.00	83.00
[49]	2021	CNN	Reddit	72.14	72.92
[63]	2020	LSTM	Reddit	65.92	65.93
[48]	2021	LSTM	Twitter	—.—	94.00
[58]	2021	LSTM	Reddit	—.—	64.00
[59]	2021	LSTM	Reddit	—.—	94.90
[61]	2022	Transformer	Reddit	—.—	79.00

of essential components such as convolutional layers, pooling layers, and fully connected layers, as illustrated in Fig. 1.

In general, LSTMs face challenges in effectively capturing very long-term dependencies. As the temporal gap between relevant elements within a sequence widens, the ability of LSTMs to retain and propagate information over extended distances diminishes. This becomes particularly problematic when dealing with variable-length sequences, such as sentences of varying lengths in NLP tasks, where processing the data efficiently with LSTMs becomes difficult. Moreover, while LSTMs can capture dependencies within a local context, they often struggle to encompass broader contextual information beyond the immediate sequence. This limitation is especially significant in tasks that require an understanding of the larger context for accurate predictions. For example, in document-level sentiment analysis, LSTMs may face difficulties incorporating information from the entire document [66,67].

In cases where long-term dependencies need to be extracted, 1D CNNs alone cannot accurately capture such dependencies through convolutional operations [68,69]. While transformers offer a powerful alternative, they come with considerable computational costs, especially when handling large models with many layers and attention heads. The self-attention mechanism, which enables every element to interact with all others in the sequence, introduces significant computational overhead. As a result, training and inference with large transformer models require substantial computational resources and time. Furthermore, despite their strength in capturing long-range dependencies, transformers sometimes struggle to fully understand contextual information beyond the local sequence. In tasks that demand a grasp of global context or document-level knowledge, transformers may require additional mechanisms or modifications to efficiently absorb and utilize this broader information [70].

Integrating diverse architectures within a hybrid model adds another layer of complexity. Effectively managing and optimizing the interactions between BiLSTM and CNN components requires careful design and training. This increased complexity can lead to challenges in model explanation, debugging, and fine-tuning. Additionally, compared to standalone architectures, hybrid models like BiLSTM-CNN tend to incur higher computational costs. The simultaneous operation of BiLSTMs and CNNs requires traversing multiple layers, which increases computational demands during training and validation. Consequently, these models may necessitate greater computational resources and time for effective training. Moreover, combining different architectures within a hybrid model may result in information loss or bottleneck effects [71,72].

Convolutional layers are fundamental building blocks of CNNs. They comprise multiple filters (also known as kernels) that slide over the input image, performing convolution operations to detect patterns. Each kernel specializes in extracting different features from the input. For instance, the 64 kernels depicted in Fig. 1 represent the first convolutional layer of a 2D pre-trained CNN model, namely AlexNet. While common CNN models typically have several convolutional layers, the accurate performance of the first layer's convolutions is crucial for correctly extracting low-level image features, such as orientation, curvature, corners, crosses, and compositions. This, in turn, enhances the performance of subsequent layers. Following the convolutional layers, pooling layers are employed to reduce the spatial dimensions of the output from the previous layers. Pooling layers downsample the data by taking the maximum or average values from groups of pixels, which helps reduce computational complexity and control overfitting. After several convolutional and pooling layers, features are extracted, and the resulting feature maps are flattened into a vector. This vector is then fed into one or more fully connected layers, which operate like traditional neural network layers, connecting all neurons from the previous layer to all neurons in the next. Finally, the values obtained from these neurons are mapped to desired classes, such as suicidal or non-suicidal.

Two-dimensional CNN models are extensively used in image classification across various applications, including object classification, animal detection, car plate recognition, handwritten digit recognition, facial identification, disease classification through medical images, and environmental monitoring via satellite image analysis. In some cases, they achieve higher accuracy than humans [73]. These models are explicitly designed to exploit hierarchical representations, in addition to the local correlations extractable by 1D CNNs, enabling them to capture both basic and local features from initial convolutions, as well as complex patterns and features from higher layers [74,75]. The incorporation of such pre-trained models provides a significant advantage for classification tasks, as they come equipped with knowledge gained from diverse image datasets, an advantage not readily available in 1D models [76–78]. Furthermore, 2D CNN models typically incorporate spatial pooling layers to aggregate information from neighboring regions, aiding in noise reduction [79]. Moreover, these models exhibit

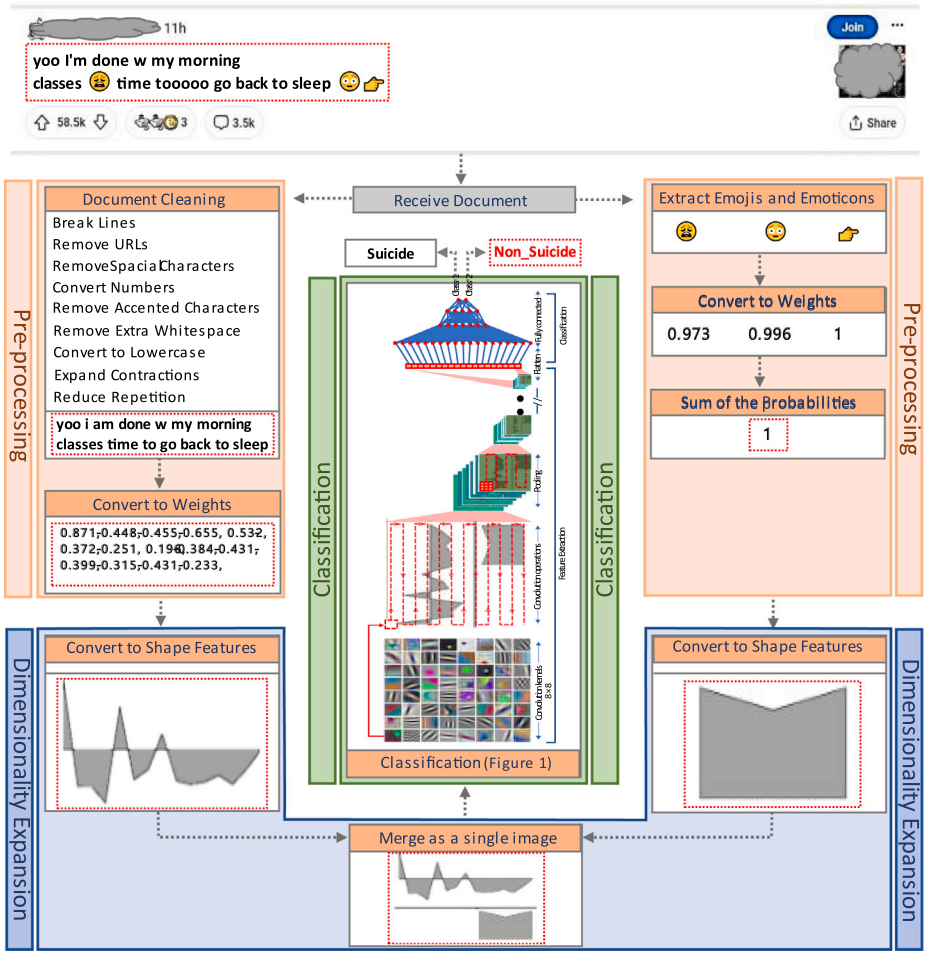

Fig. 2. The proposed model for informal document classification uses dimensionality expansion. In this approach, texts and signs are separated from each other (e.g., in suicide and non-suicide documents). Data cleaning operations are applied to the texts, and both the texts and signs are converted into suitable images. These images are then merged and fed into a CNN model for classification.

a degree of translation invariance, allowing them to recognize objects or patterns regardless of their position within the image. This characteristic enhances the model's robustness to minor spatial variations or noise-induced shifts in the data [80].

Another advantage of 2D CNN models is the use of transfer learning. Training deep learning models from scratch can be computationally intensive and resource-consuming. Utilizing pre-trained models and transfer learning can significantly improve the efficiency of the training process. VGG-16, ResNet-50, and AlexNet are three common examples of 2D pre-trained CNN models, all trained on the ImageNet dataset, which contains 14 million images. In the remainder of this article, we employ these three models as our subjects of investigation.

4. Proposed model

The proposed model for informal document classification of suicide and non-suicide posts is illustrated in Fig. 2. As shown in the figure, the model comprises three main stages: document preprocessing, dimensionality expansion, and classification. First, the formal or informal document, containing both text and symbols, is processed. Symbols are separated from the text, and after several stages of cleaning, both words and symbols are transformed into numerical representations based on their weight in different classes. These numerical weights are then converted into visual shapes (images) using a visual concept similar to an area plot. The resulting images for each document are fused into a single image, which is then fed into a pre-trained 2D CNN model. The following sections provide a detailed explanation of each step.

4.1. Data pre-processing operations

Data pre-processing is a crucial step in document classification tasks. It involves cleaning, converting, and organizing raw text data into a format suitable for machine learning algorithms. Effective pre-processing enhances the quality of the input data and improves the performance of the classification model. Models that receive 1D data typically analyze documents as pure text, although these documents often contain not only text but also other elements such as emojis and emoticons, which are commonly removed during data cleaning. However, these symbols are often used to convey opinions and emotions. Therefore, correctly accounting for these symbols in the analysis can significantly increase model accuracy. In our approach, we extract these symbols from the documents during the pre-processing stage, prior to performing traditional data cleaning.

For example, Fig. 3 illustrates the frequency of emoji occurrences in a sample Reddit dataset related to suicide and non-suicide content. The horizontal axis shows the sequence of emojis most frequently used in the dataset, while the vertical axis shows the number of occurrences for each emoji in both classes. In nearly all cases, the frequency of emoji usage in the non-suicide class is higher than in the suicide class. The appearance of the e_i emoji in a document suggests that the document is likely not related to suicide, with a probability of $P\left(e_i\left(non\right)\right)$. This probability, $P\left(e_i\left(non\right)\right)$, is calculated for each document based on Eq. (1), where $N\left(e_i\left(non\right)\right)$ represents the number of occurrences of a specific emoji in the non-suicide class, $N\left(e_i\left(T\right)\right)$ represents the total number of occurrences of that emoji in both the suicide and non-suicide classes, and n denotes the total number of emoji types in the dataset.

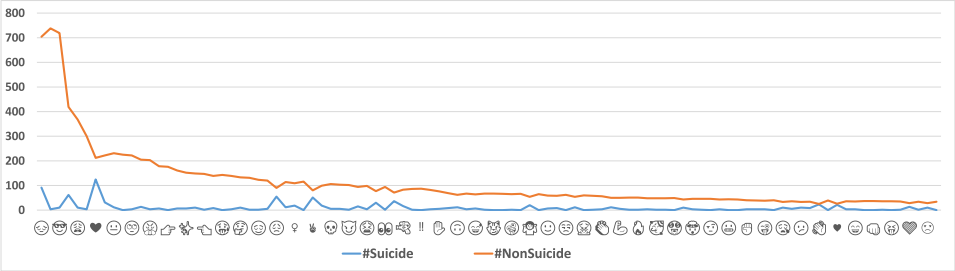

Fig. 3. Some examples of emojis extracted from the Reddit social network, along with their frequency in both the suicide and non-suicide classes, sorted from highest to lowest repetition.

$$P\left(e_i\left(non\right)\right) = \frac{N\left(e_i\left(non\right)\right)}{N\left(e_i\left(T\right)\right)} \qquad 1 \le i \le n \tag{1}$$

If several different signs appear in a document, according to Eq. (2), the probability that a document belongs to non-suicide (P_D), is obtained from the sum rule of the independent probabilities of every sign that appeared in that document.

$$P_D = 1 - \left[\left(1 - P\left(e_1(\text{non})\right)\right) \times \left(1 - P\left(e_2(\text{non})\right)\right) \times \cdots \times \left(1 - P\left(e_n(\text{non})\right)\right)\right] \tag{2}$$

In some applications, such as user-generated posts on social networks, informal language is used, which often contains various grammatical and lexical errors, as well as additional signs and symbols that can hinder the accurate analysis of documents by deep learning models. During the preprocessing stage, after separating signs from the documents, we aimed to mitigate the impact of these factors by performing several data-cleaning operations. These included removing line breaks, URLs, special characters, and accented characters; converting text to lowercase; expanding contractions; and reducing repetitions.

Next, each word in each class is treated as an object, and the difference in its frequency across different classes is calculated for the entire document. In Eq. (3), W_{w_i} represents the calculation of the weight of each word (object), where $N(w_i(\text{non}))$ and $N(w_i(\text{sui}))$ are the number of occurrences of each word in the non-suicide and suicide classes, respectively, and $N(W_i(\text{T}))$ is the normalized word count across the entire dataset. Since the number of documents in the non-suicide class is 1.66 times that of the suicide class, a coefficient of 1.66 is used to normalize the weight ratio of words between the two classes. In this case, the greater the weight of an object in the non-suicide class compared to the suicide class, the value will be between 0 and 1. Conversely, if the weight in the suicide class is greater than in the non-suicide class, the value will be between 0 and −1.

$$W_{w_i} = \frac{N(w_i(\text{non})) - (N(w_i(\text{sui})) \times 1.66)}{N(W_i(\text{T}))} \tag{3}$$

For example, consider the sentence: "*I need help, just help me girls. I am crying so hard.*" In Table 2, the frequency of each object in the *Suicide* and *Non-suicide* classes is specified.

For the object *crying*, its occurrence in the *Suicide* class is 1417, and in the *Non-suicide* class, it is 723. Substituting these values into Eq. (3), the weight W_{crying} is calculated as follows:

$$W_{\text{crying}} = \frac{723 - (1417 \times 1.66)}{3057} \approx -0.53$$

The value 3057 represents the total number of observations of the *crying* object in a normalized state across the entire dataset. Since the weight of the object *crying* is negative, it indicates that this object is a significant factor in suicide-related content.

Similar steps are taken for the preprocessing of signs. In the next step, the numbers corresponding to the words and signs in the documents were converted into images.

Table 2
Word frequencies associated with the sentence *"I need help, just help me. I am crying so hard."* F.Suicide represents the frequency of a word (object) in the suicide class, while F.Non-suicide represents the frequency in the non-suicide class.

Num.	Word	F.Suicide	F.Non-suicide	W_{w_i}
1	i	405 341	256 628	−0.45
2	need	8930	7958	−0.30
3	help	12 133	7556	−0.45
4	just	50 665	34 984	−0.41
5	help	12 133	7556	−0.45
6	me	61 506	45 473	−0.38
7	girls	340	3452	+0.72
8	i	405 341	256 628	−0.45
9	am	99 618	61 942	−0.46
10	crying	1417	723	−0.53
11	so	35 523	36 253	−0.24
12	hard	3457	2099	−0.46

4.2. Dimensionality expansion

In dimensionality expansion, data and patterns are transformed from a lower-dimensional space to a higher-dimensional one. To improve the accuracy of binary document classification, we changed the dimensionality of the data from 1D to 2D images. This approach eliminates the need to design and train new models from scratch, allows features to be visualized, and preserves privacy through the transformation. In this dimensionality expansion stage, the weights produced during preprocessing are converted into shapes. The shapes of signs and text related to each document are merged to form a single image.

As explained in [81], images with simple geometric patterns, such as curves, corners, crosses, and edges, can reduce the cost of fine-tuning. This is because the early convolution layers of 2D pre-trained CNN models can effectively extract these features. To leverage this, we used solid corners, which fall under the category of simple shapes that can be extracted by the initial layers of various two-dimensional CNN models, to convert one-dimensional values into two-dimensional ones. Fig. 4 demonstrates how different models are able to extract solid corners in their first convolutional layers. To compare kernels with similar performance, each one was evaluated against Gabor filters, which are used to extract corners, using the Structural Similarity Index (SSIM) method. The Gabor transform can be fine-tuned by adjusting the sinusoidal wave parameters (amplitude, frequency, and phase) and the Gaussian envelope (standard deviation). Gabor filters are generated in various sizes, orientations, and frequencies to extract solid corners. Kernels with an SSIM score above 80% were considered effective for extracting solid corners. As shown in the figure, all three pre-trained models – VGG-16, ResNet-50, and AlexNet – are capable of extracting solid corners.

Converting values into images based on the capabilities of the first convolution layers leads to a reduction in the cost of fine-tuning. For example, the changes of the kernels within the first convolution layer of the AlexNet model are shown in Fig. 5. Kernels were compared one

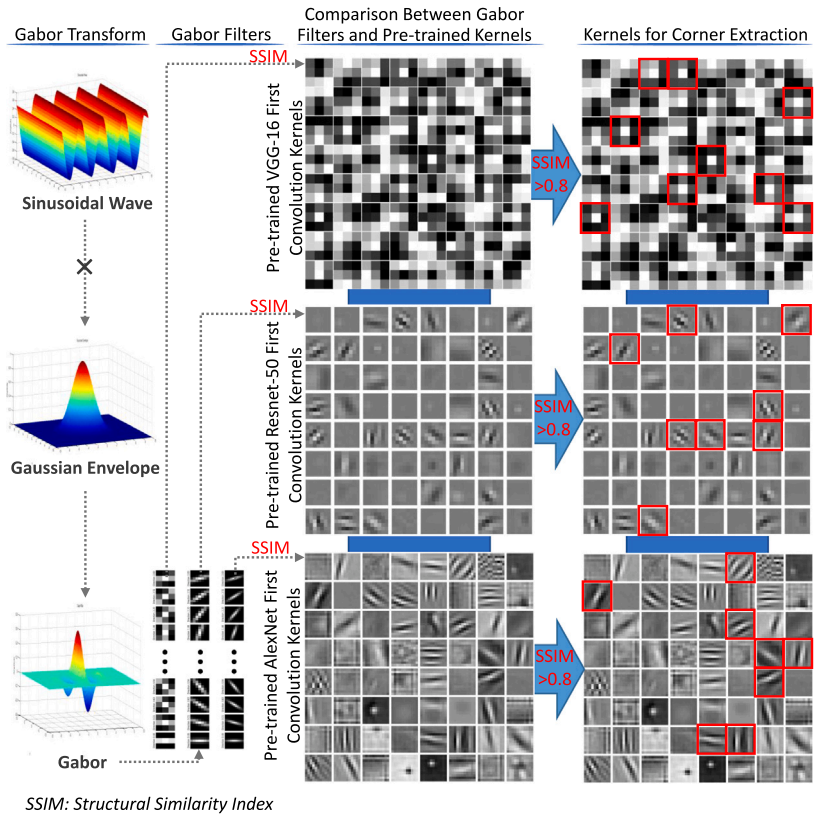

Fig. 4. Selecting kernels that are able to extract solid corners. Different Gabor filters for extracting solid corners are compared with the kernels of pre-trained models using the SSIM method. Kernels with a similarity greater than 80% are marked with a red border.

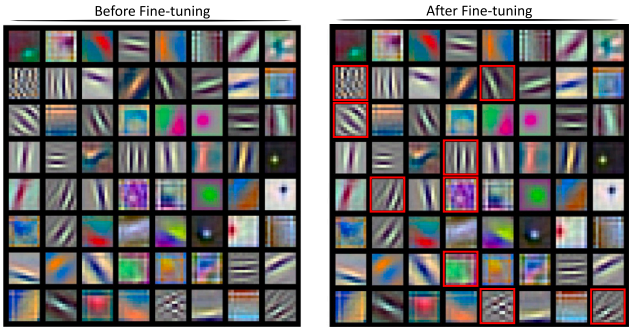

Fig. 5. Comparison between first convolution layer kernels of the 2D pre-trained AlexNet model, before and after fine-tuning. Only nine kernels have changed more than 1% and are marked with red border.

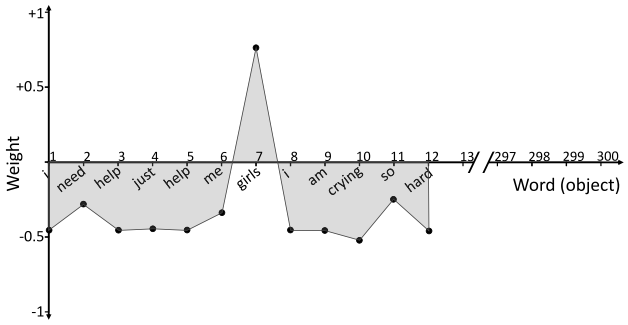

Fig. 6. The conversion of words in a sentence into weights, followed by their transformation into a two-dimensional representational structure. The horizontal axis represents the words (objects), and the vertical axis represents the normalized weight for each object. Values greater than zero indicate an object's tendency towards non-suicide, while values less than zero indicate an object's tendency towards suicide.

by one with the help of the SSIM method. Nine kernels marked with a red border have been changed by ranging from 1% to 3%. Most kernels have been changed less than 1% after a hundred epochs. This means that the fine-tuning operation using the created images has little effect on the weights.

We converted the calculated weights related to texts to angular gray shapes (corners), which can be seen in Fig. 6.

To produce an image for a document, horizontal and vertical axes are used to represent the number of words in a document and the weight of a word in that document, respectively. A horizontal line with a length of m pixels can represent a maximum of m words. In order to reduce the negative impact of the resizing operations for matching the input image to the 2D pre-trained CNN models, 300 pixels have been considered. This value is sufficient for some document classification applications such as mental disorder detection. The values of the calculated weights range from −1 to +1. A shape is formed by joining the points to each other. The area of the created shape has been highlighted by grayscale in our figures. Each weight in each created shape from left to right is proportional to the appearance of a word in order in a document from top to down and left to right. In other words, the order of occurrence is automatically preserved in the model training process. The shape that represents the total weight of emojis from −1 jo 1 in each document consists of two solid corners that are connected to each other and create a two-sided sloping surface. In this way, the change in the sloping surface causes a change in the area below it, so that a lower slope indicates more emoji weight. In the next step, to fuse signs with texts, two created images, one for text and another for signs are merged. The final created image represents a document that is used as an input image in the classification stage. Fig. 7 shows how text and related symbols from social media posts are fused into single images,

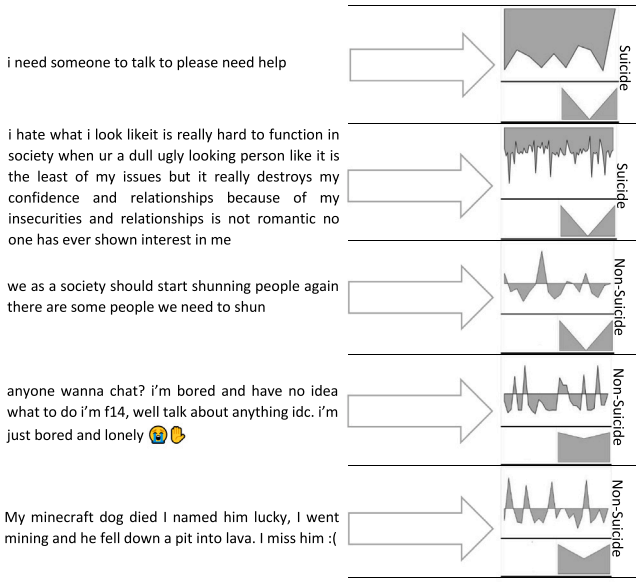

Fig. 7. Examples of fusing text and corresponding symbols from a post into a single image for both suicide and non-suicide posts. The shape corresponding to each text is placed in the upper half of the image, while the shape corresponding to the total weight of emoji and emoticon probabilities is placed in the lower right half of the image.

distinguishing between posts related to suicide and those that are not. The top part of each image displays the visual representation of the text, while the bottom right shows the calculated significance of emojis and emoticons. By performing this whole operation, the documents data related to the "suicide and depression detection" dataset [82] are converted to images.

4.3. Classification

In the classification stage, the fused textual and emoji-based images are fed into 2D pre-trained CNN models. For our classification task, we employed three widely used models—VGG-16, ResNet-50, and AlexNet—all of which were pre-trained on the ImageNet dataset. These models are chosen for their proven ability to generalize well to new image-based tasks, allowing us to leverage transfer learning effectively.

Each model was fine-tuned on 167,290 images derived from suicide and non-suicide documents, specifically structured for the binary classification task of identifying suicide-related content. Fine-tuning involved adjusting the higher layers of these pre-trained networks to better recognize subtle patterns unique to this dataset, such as specific textual cues and emotive symbols indicative of suicidal ideation. The choice of CNN models offers distinct advantages: VGG-16 provides a deep, sequential layer structure that is efficient for capturing hierarchical patterns; ResNet-50 incorporates residual connections that prevent the vanishing gradient problem, enabling better performance on deeper networks; and AlexNet, with its comparatively simpler architecture, serves as a baseline for rapid prototyping and validation.

Our experimental results demonstrated that these models, particularly when fine-tuned, achieved remarkably high accuracy. The VGG-16 model, for instance, reached an accuracy of 99.99%, underscoring the effectiveness of dimensionality expansion and the robustness of 2D CNNs in identifying complex patterns within transformed 1D textual data. This accuracy highlights the potential of our approach to contribute to real-world applications where precise classification of mental health-related content is crucial.

5. Experimental results

5.1. Experimental setup

Data cleaning operation was performed by macro commands in Excel and also in the Jupiter notebook environment. Dimensionality expansion operation was performed by Visual Basic and finally, training and classification operation was performed in Colab environment by Keras and PyTorch. The hardware configuration included an NVIDIA Tesla P100 GPU, 87 GB of RAM, and 180 GB of storage space. The "Suicide and Depression Detection" dataset is developed from self-reported posts on the social media platform Reddit, and includes text, emoji and emoticons with a length of one to more than 140 tokens. The dataset divides posts into two main categories: non-suicide, with approximately 169,518 samples, and suicide, with around 63,721 samples. In the suicidal category, users have indicated that they either had thoughts of suicide or made attempts following their posts. The data, which has been collected since 2018, provides valuable insights into the language patterns associated with depressive and suicidal ideation, making it a crucial resource for training machine learning models aimed at detecting mental health crises. In the dimensionality expansion stage, in order to remove the resizing operations in the 2D CNN models, images are created in the size of 300×300. Three 2D CNN models VGG-16, ResNet-50, and AlexNet are investigated. Due to the similarity of the obtained results for the three models, the results related to VGG-16 are presented and discussed.

5.2. Evaluation results

Converting a document consisting of textual data and signs into images and employing 2D pre-trained CNN models, which possess the capability to extract features at both local and global levels, leads to an improvement in classification accuracy compared to baseline models. Table 3 shows this comparison using Alexnet, Resnet-50, and VGG-16 CNNs as 2D pre-trained models on the Reddit dataset [82], which has been transformed into images using the proposed dimensionality expansion technique. As shown in the table, the accuracy (second column) of the baseline models ranges from 58.7% to 97.40%, while the accuracy of the proposed model reaches an impressive 99.99% for VGG-16.

In order to show the impact of fine-tuning on the 2D models, we have used 2D CNN models in two cases, with pre-training and without pre-training. In other words, we have applied our proposed model on Alexnet, Resnet-50, and VGG-16 CNNs with pre-trained and without pre-trained. Additionally, it can be inferred that while the selection of the CNN model type impacts the model's accuracy by 2 to 4 percent in various conditions, the presence or absence of symbols has a more significant effect on accuracy. Fig. 8 depicts the difference in validation accuracy with and without using ImageNet weights. In fact, the pre-trained VGG-16 CNN model achieves the highest accuracy in the 5th epoch, while without pre-trained, it achieves the highest accuracy in the 100th epoch. This means that the computational cost is reduced 20 times by 2D pre-trained models. We can use the weights and patterns learned by the pre-trained 2D CNN models as a starting point, and fine-tune them to our specific task. Converting documents to simple images that contain simple geometric shapes helped to reduce fine-tuning costs.

According to Table 3 in order to support the results, experiments are tested on the "Reddit Self-reported Depression Diagnosis" (RSDD) dataset, which contains documents related to depression diagnoses. Additionally, the impact of the presence or absence of symbols on model performance was examined. In the pre-trained models, the presence of symbols improved model accuracy by approximately 8 percent.

Table 3
Comparison between baseline models and the 2D pre-trained VGG-16 model in document classification of the *"Self-reported Depression Diagnosis"* dataset, which comprises both depressed and non-depressed documents. The term "Pre-trai.n-s" refers to cases where symbols, such as emojis, were not considered by the model, while "Pre-trai." refers to cases where these symbols were included in the model's analysis.

Model name	Acc. (%)	Pre. (%)	Rec. (%)
1D-CNN	58.17	62.90	65.75
LSTM	63.65	67.35	70.05
BiLSTM-CNN	69.19	69.83	69.90
Chat GPT-3	79.74	81.73	82.88
RoBERTa	83.72	84.94	86.91
Ref. [83]	88.48	87.36	94.57
Ref. [57]	90.30	91.06	93.70
Ref. [53]	95.00	94.30	94.90
Ref. [64]	97.40	99.92	97.40
Non pre-trained 2D models without considering symbols			
Non pre-trai.n-s AlexNet	86.15	86.27	86.73
Non pre-trai.n-s Resnet-50	86.75	86.82	87.09
Non pre-trai.n-s VGG-16	87.04	88.22	89.49
Non pre-trained 2D models, considering symbols			
Non pre-trai. AlexNet	94.38	94.68	94.96
Non pre-trai. Resnet-50	95.27	97.59	98.45
Non pre-trai. VGG-16	98.51	98.73	98.05
Pre-trained 2D models, without considering symbols			
Pre-trai.n-s AlexNet	90.81%	91.04%	90.96%
Pre-trai.n-s Resnet-50	92.01%	91.93%	91.90%
Pre-trai.n-s VGG-16	93.71%	94.23%	93.84%
Pre-trained 2D models, considering symbols			
Pre-trai. AlexNet	97.89	97.90	97.73
Pre-trai. Resnet-50	99.67	99.64	99.49
Pre-trai. VGG-16	**99.99**	**99.99**	**99.99**

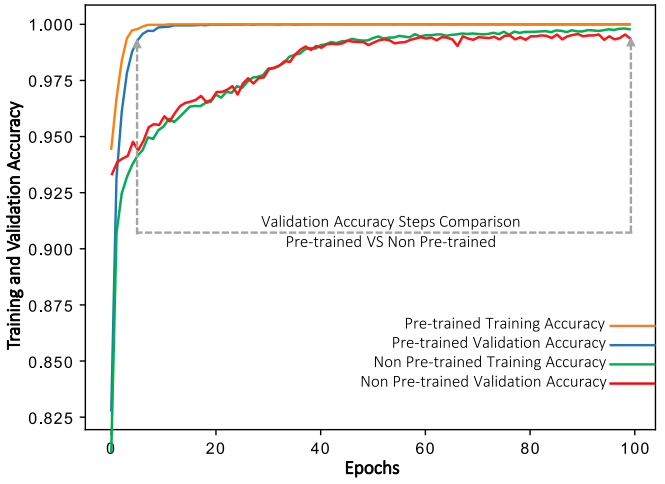

Fig. 8. Impact of fine-tuning operation on the VGG-16 model with pre-trained and without pre-trained. In the pre-trained mode which 2D CNN uses ImageNet weights achieves the highest validation accuracy in the 5th epoch, while training the VGG-16 model from scratch achieves the highest validation accuracy in the 100th epoch.

5.3. Discussion

Fig. 9 depicts two examples of documents related to non-suicide (top) and suicide (bottom) and their corresponding created images on the right side. The color map on each document is proportional to its weight in the corresponding image. In the figure, explainability is added to the model so that it can be concluded which word has a greater impact on the document's tendency towards suicidal or non-suicidal. If there is an sign in the document, its weight is placed separately in the image. The gauge on the right side is the ratio of the area of the upper part of the horizontal axis (non-suicide) to the lower part of it (suicide). From this ratio, it is not possible to definitively conclude which class the document belongs to, but it shows its tendency towards each of the classes.

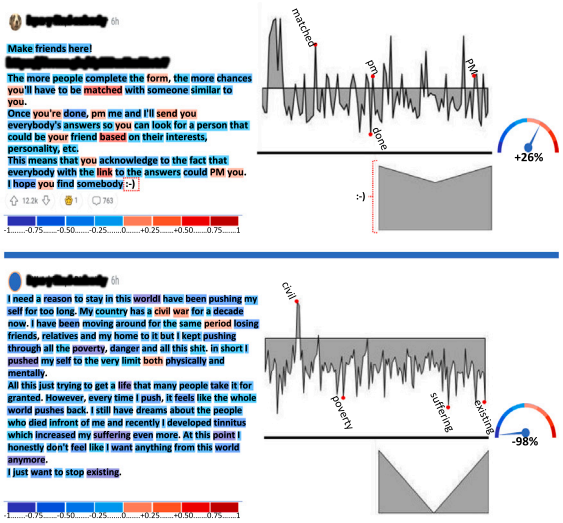

Fig. 9. Comparison between converting non-suicide (top) and suicide (bottom) documents to images. On the right side of each document, the corresponding created image is shown. The color map in each document corresponds to the weight of each word. The gauge shown on the right side shows the overall weight of the document/image towards suicide or non-suicide.

Instead of transferring and displaying texts in various applications related to informal document-level text classification, where personal data may be leaked and revealed to unauthorized people, images related to documents are used, which, like a security layer, hides words from the view of unauthorized people.

In the sample dataset that we used, the number of words in each document was less than 300. If the word count of a document is more than 300 words, a sliding window with dimensions of 300 words is considered. For each sliding window, a new shape is created. As shown in Fig. 10 the number of shapes in each generated image is limited to 14 which means a created image includes up to 4200 words. For a document with a size larger than 4200 words, different images (frames) will be created and the created images will be considered as a 3D representation of that document.

By converting documents in the "*suicide and depression detection*" dataset to images, an image dataset containing 167,290 images across two classes is created. One class consists of suicide-related content with 62,713 image samples, while the other class contains non-suicidal content with 104,577 images.

6. Conclusions

This study has developed a novel approach for detecting suicide-related documents by converting 1D informal documents containing text and symbols into 2D images through a dimensionality expansion technique. By leveraging pre-trained 2D CNN models, including AlexNet, ResNet-50, and VGG-16, our method achieved an impressive classification accuracy of 99.99%, outperforming traditional baseline models. This technique not only enhanced classification accuracy but also significantly reduced computational costs by utilizing the pre-existing learned features of the CNN models, an efficiency that is crucial for practical applications where resource optimization is essential. Furthermore, transforming textual data into visual representations allows for the integration of both textual and visual cues, providing a comprehensive analysis that captures subtle nuances in documents. This dual-modal analysis is particularly effective in sensitive applications like suicide detection, where accuracy and early detection are

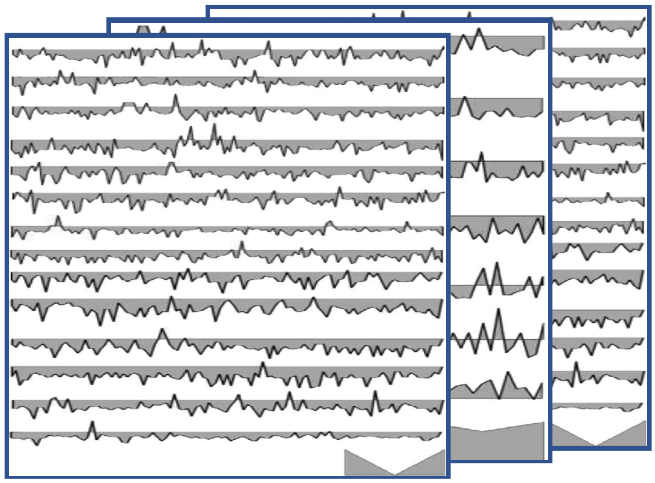

Fig. 10. Representing documents with more than 300 words, each sliding window contains 300 words for each shape per line in the output image. The final output image contains up to 4200 words. Documents with more than 4200 words are displayed on multiple images (frames).

paramount. We also addressed the ethical implications and prioritized data privacy in handling sensitive information. Additionally, this visual representation serves as a complementary method to explain the results. Overall, this research underscores the potential of dimensionality expansion, combined with advanced deep learning models, to address complex classification tasks across various domains.

CRediT authorship contribution statement

Nima Esmi: Writing – original draft, Visualization, Project administration, Methodology, Investigation. **Asadollah Shahbahrami:** Writing – review & editing. **Georgi Gaydadjiev:** Writing – review & editing. **Peter de Jonge:** Writing – review & editing.

Ethics statement

This study utilized the publicly available "Suicide and Depression Detection" dataset from Reddit, consisting of self-reported social media posts. As the research involved secondary analysis of pre-existing, anonymized data and did not entail direct experimentation with human subjects, no formal ethical approval or informed consent was required. The dataset was handled in compliance with relevant data privacy laws and institutional guidelines at the University of Groningen, ensuring that no personally identifiable information was accessed or disclosed. The dimensionality expansion technique further obscured individual data, enhancing privacy protection. Thus, an Ethics statement with approval details is not applicable to this study.

Declaration of competing interest

None declared.

References

[1] E. Hossain, R. Rana, N. Higgins, J. Soar, P.D. Barua, A.R. Pisani, K. Turner, Natural language processing in electronic health records in relation to healthcare decision-making: a systematic review, Comput. Biol. Med. 155 (2023) 106649, http://dx.doi.org/10.1016/j.compbiomed.2023.106649.

[2] R. Chiong, G.S. Budhi, S. Dhakal, F. Chiong, A textual-based featuring approach for depression detection using machine learning classifiers and social media texts, Comput. Biol. Med. 135 (2021) 104499, http://dx.doi.org/10.1016/j.compbiomed.2021.104499.

[3] L. Samaras, E. García-Barriocanal, M.-A. Sicilia, Sentiment analysis of covid-19 cases in greece using twitter data, Expert Syst. Appl. 230 (2023) 1–9, http://dx.doi.org/10.1016/j.eswa.2023.120577.

[4] W. Zhang, X. Li, Y. Deng, L. Bing, W. Lam, A survey on aspect-based sentiment analysis: Tasks, methods, and challenges, IEEE Trans. Knowl. Data Eng. 35 (11) (2022) 11019–11038, http://dx.doi.org/10.1109/TKDE.2022.3230975.

[5] S. Rao, A.K. Verma, T. Bhatia, Hybrid ensemble framework with self-attention mechanism for social spam detection on imbalanced data, Expert Syst. Appl. 217 (2023) 1–21, http://dx.doi.org/10.1016/j.eswa.2023.119594.

[6] X. Ma, J. Wu, S. Xue, J. Yang, C. Zhou, Q.Z. Sheng, H. Xiong, L. Akoglu, A comprehensive survey on graph anomaly detection with deep learning, IEEE Trans. Knowl. Data Eng. 35 (12) (2021) 12012–12038, http://dx.doi.org/10.1109/TKDE.2021.3118815.

[7] S.B. Abkenar, M.H. Kashani, M. Akbari, E. Mahdipour, Learning textual features for twitter spam detection: A systematic literature review, Expert Syst. Appl. 228 (2023) 1–27, http://dx.doi.org/10.1016/j.eswa.2023.120366.

[8] B. Han, I.W. Tsang, X. Xiao, L. Chen, S.-F. Fung, C.P. Yu, Privacy-preserving stochastic gradual learning, IEEE Trans. Knowl. Data Eng. 33 (8) (2021) 3129–3140, http://dx.doi.org/10.1109/TKDE.2020.2963977.

[9] G. Cola, M. Mazza, M. Tesconi, Twitter newcomers: Uncovering the behavior and fate of new accounts through early detection and monitoring, IEEE Access 11 (2023) 55223–55232, http://dx.doi.org/10.1109/ACCESS.2023.3282580.

[10] M. Mazza, M. Avvenuti, S. Cresci, M. Tesconi, Investigating the difference between trolls, social bots, and humans on twitter, Comput. Commun. 196 (2022) 23–36, http://dx.doi.org/10.1016/j.comcom.2022.09.022.

[11] S. Aftan, H. Shah, Using the arabert modhael for customer satisfaction classification of telecom sectors in Saudi Arabia, Brain Sci. 13 (1) (2023) 1–20, http://dx.doi.org/10.3390/brainsci13010147.

[12] L. Hu, S. Wei, Z. Zhao, B. Wu, Deep learning for fake news detection: A comprehensive survey, AI Open 3 (2022) 133–155, http://dx.doi.org/10.1016/j.aiopen.2022.09.001.

[13] B.-Y. Hsu, C.-Y. Shen, X. Yan, Network intervention for mental disorders with minimum small dense subgroups, IEEE Trans. Knowl. Data Eng. 33 (5) (2021) 2121–2136, http://dx.doi.org/10.1109/TKDE.2019.2949294.

[14] M. Kabir, T. Ahmed, M.B. Hasan, M.T.R. Laskar, T.K. Joarder, H. Mahmud, K. Hasan, Deptweet: A typology for social media texts to detect depression severities, Comput. Hum. Behav. 139 (2023) 1–14, http://dx.doi.org/10.1016/j.chb.2022.107503.

[15] M. Khorasani, M. Kahani, S.A.A. Yazdi, M. Hajiaghaei-Keshteli, Towards finding the lost generation of autistic adults: A deep and multi-view learning approach on social media, Knowl.-Based Syst. 216 (2023) 1–17, http://dx.doi.org/10.1016/j.knosys.2023.110724.

[16] O. Edo-Osagie, B. De La Iglesia, I. Lake, O. Edeghere, Ascoping review of the use of twitter for public health research, Comput. Biol. Med. 122 (2020) 103770, http://dx.doi.org/10.1016/j.compbiomed.2020.103770.

[17] W. Ragheb, J. Aze, S. Bringay, M. Servajean, Negatively correlated noisy learners for at-risk user detection on social networks: A study on depression, anorexia, self-harm and suicide, IEEE Trans. Knowl. Data Eng. 35 (1) (2021) 770–783, http://dx.doi.org/10.1109/TKDE.2021.3078898.

[18] A.H. Saffar, T.K. Mann, B. Ofoghi, Textual emotion detection in health: Advances and applications, J. Biomed. Inf. 137 (2022) 1–15, http://dx.doi.org/10.1016/j.jbi.2022.104258.

[19] M. Garg, Mental health analysis in social media posts: a survey, Arch. Comput. Methods Eng. 30 (3) (2023) 1819–1842, http://dx.doi.org/10.1007/s11831-022-09863-z.

[20] L.Y. Jiang, X.C. Liu, N.P. Nejatian, M. Nasir-Moin, D. Wang, A. Abidin, K. Eaton, H.A. Riina, I. Laufer, P. Punjabi, et al., Health system-scale language models are all-purpose prediction engines, Nature 619 (2023) 357–362, http://dx.doi.org/10.1038/s41586-023-06160-y.

[21] V. Mhasawade, Y. Zhao, R. Chunara, Machine learning and algorithmic fairness in public and population health, Nat. Mach. Intell. 3 (8) (2021) 659–666, http://dx.doi.org/10.1038/s42256-021-00373-4.

[22] W.H. Kwok, Y. Zhang, G. Wang, Artificial intelligence in perinatal mental health research: A scoping review, Comput. Biol. Med. (2024) 108685, http://dx.doi.org/10.1016/j.compbiomed.2024.108685.

[23] S. Delgado, R.C.B. Vignola, R.J. Sassi, P.A. Belan, S.A. de Araújo, Symptom mapping and personalized care for depression, anxiety and stress: A data-driven ai approach, Comput. Biol. Med. 182 (2024) 109146, http://dx.doi.org/10.1016/j.compbiomed.2024.109146.

[24] S. Selva Birunda, R. Kanniga Devi, Areview on word embedding techniques for text classification, in: Innovative Data Communication Technologies and Application: Proceedings of ICIDCA 2020, 2021, pp. 267–281, http://dx.doi.org/10.1007/978-981-15-9651-3_23.

[25] B. Yin, F. Corradi, S.M. Bohté, Accurate and efficient time-domain classification with adaptive spiking recurrent neural networks, Nat. Mach. Intell. 3 (10) (2021) 905–913, http://dx.doi.org/10.1038/s42256-021-00397-w.

[26] L. Hong, J. Lin, S. Li, F. Wan, H. Yang, T. Jiang, D. Zhao, J. Zeng, Anovel machine learning framework for automated biomedical relation extraction from large-scale literature repositories, Nat. Mach. Intell. 2 (6) (2020) 347–355, http://dx.doi.org/10.1038/s42256-020-0189-y.

[27] S. Kench, S.J. Cooper, Generating three-dimensional structures from a two-dimensional slice with generative adversarial network-based dimensionality expansion, Nat. Mach. Intell. 3 (4) (2021) 299–305, http://dx.doi.org/10.1038/s42256-021-00322-1.

[28] Q. Yang, W. Jin, Q. Zhang, Y. Wei, Z. Guo, X. Li, Y. Yang, Q. Luo, H. Tian, T.-L. Ren, Mixed-modality speech recognition and interaction using a wearable artificial throat, Nat. Mach. Intell. 5 (2) (2023) 169–180, http://dx.doi.org/10.1038/s42256-023-00616-6.

[29] Y.-C. Tsai, J.-H. Chen, Enhancing model explainability in financial trading using training aid samples: A cnn-based candlestick pattern recognition approach, in: International Conference on Information Reuse and Integration for Data Science, 2023, pp. 76–82, http://dx.doi.org/10.1109/IRI58017.2023.00021.

[30] M. Mahmood, D. Mzurikwao, Y.-S. Kim, Y. Lee, S. Mishra, R. Herbert, A. Duarte, C.S. Ang, W.-H. Yeo, Fully portable and wireless universal brain–machine interfaces enabled by flexible scalp electronics and deep learning algorithm, Nat. Mach. Intell. 1 (9) (2019) 412–422, http://dx.doi.org/10.1038/s42256-019-0091-7.

[31] X. Xiong, T. Zhu, Y. Zhu, M. Cao, J. Xiao, L. Li, F. Wang, C. Fan, H. Pei, Molecular convolutional neural networks with dna regulatory circuits, Nat. Mach. Intell. 4 (7) (2022) 625–635, http://dx.doi.org/10.1038/s42256-022-00502-7.

[32] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, Adv. Neural Inf. Process. Syst. 25 (2012) 1–9, http://dx.doi.org/10.1145/3065386.

[33] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Computer Vision and Pattern Recognition, 2016, pp. 770–778, http://dx.doi.org/10.48550/arXiv.1512.03385.

[34] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, 2014, http://dx.doi.org/10.48550/arXiv.1409.1556, arXiv preprint arXiv:1409.1556.

[35] M. Jusup, P. Holme, K. Kanazawa, M. Takayasu, I. Romić, Z. Wang, S. Geček, T. Lipić, B. Podobnik, L. Wang, et al., Social physics, Phys. Rep. 948 (2022) 1–148, http://dx.doi.org/10.1016/j.physrep.2021.10.005.

[36] K. Kumari, J.P. Singh, Multi-modal cyber-aggression detection with feature optimization by firefly algorithm, Multimed. Syst. 18 (1) (2021) 1–12, http://dx.doi.org/10.1007/s00530-021-00785-7.

[37] S. Bharti, A.K. Yadav, M. Kumar, D. Yadav, Cyberbullying detection from tweets using deep learning, Kybernetes (2021) http://dx.doi.org/10.1108/K-01-2021-0061.

[38] S. Pericherla, E. Ilavarasan, Astudy of machine learning approaches to detect cyberbullying, in: Communication Software and Networks, Springer, Singapore, 2021, pp. 369–377, http://dx.doi.org/10.1007/978-981-15-5397-4_38.

[39] M.M. Islam, M.A. Uddin, L. Islam, A. Akter, S. Sharmin, U.K. Acharjee, Cyberbullying detection on social networks using machine learning approaches, in: Asia-Pacific Conference on Computer Science and Data Engineering, 2020, pp. 1–6, http://dx.doi.org/10.1109/CSDE50874.2020.9411601.

[40] Y. Tao, M. Yang, H. Li, Y. Wu, B. Hu, Depmstat: Multimodal spatio-temporal attentional transformer for depression detection, IEEE Trans. Knowl. Data Eng. 36 (7) (2024) 2956–2966, http://dx.doi.org/10.1109/TKDE.2024.3350071.

[41] M. Trotzek, S. Koitka, C.M. Friedrich, Utilizing neural networks and linguistic metadata for early detection of depression indications in text sequences, IEEE Trans. Knowl. Data Eng. 32 (3) (2018) 588–601, http://dx.doi.org/10.1109/TKDE.2018.2885515.

[42] T. Niederkrotenthaler, B. Till, N.D. Kapusta, M. Voracek, K. Dervic, G. Sonneck, Copycat effects after media reports on suicide: A population-based ecologic study, Soc. Sci. Med. 69 (7) (2009) 1085–1090, http://dx.doi.org/10.1016/j.socscimed.2009.07.041.

[43] B. Desmet, V. Hoste, Emotion detection in suicide notes, Expert Syst. Appl. 40 (16) (2013) 6351–6358, http://dx.doi.org/10.1016/j.eswa.2013.05.050.

[44] J. Jashinsky, S.H. Burton, C.L. Hanson, J. West, C. Giraud-Carrier, M.D. Barnes, T. Argyle, Tracking suicide risk factors through twitter in the us, Crisis 35 (1) (2014) 1–9, http://dx.doi.org/10.1027/0227-5910/a000234.

[45] H. Sueki, The association of suicide-related twitter use with suicidal behaviour: a cross-sectional study of young internet users in japan, J. Affect. Disord. 170 (2015) 155–160, http://dx.doi.org/10.1016/j.jad.2014.08.047.

[46] I. Hossen, T. Islam, M. Rashed, D. Das, et al., Early suicide prevention: Depression level prediction using machine learning and deep learning techniques for bangladeshi facebook users, in: International Conference on Fourth Industrial Revolution and beyond, 2022, pp. 735–747.

[47] A. Chadha, B. Kaushik, Performance evaluation of learning models for identification of suicidal thoughts, Comput. J. 65 (1) (2022) 139–154, http://dx.doi.org/10.1093/comjnl/bxab060.

[48] D. Bhattacharya, A. Shahina, et al., Early detection of suicidal tendencies from text data using lstm, in: Innovations in Power and Advanced Computing Technologies, 2021, pp. 1–5, http://dx.doi.org/10.1109/i-PACT52855.2021.9696630.

[49] A. Haque, V. Reddi, T. Giallanza, Deep learning for suicide and depression identification with unsupervised label correction, in: Artificial Neural Networks and Machine Learning, 2021, pp. 436–447, http://dx.doi.org/10.1007/978-3-030-86383-8_35.

[50] H. Yao, S. Rashidian, X. Dong, H. Duanmu, R.N. Rosenthal, F. Wang, et al., Detection of suicidality among opioid users on reddit: Machine learning–based approach, J. Med. Internet Res. 22 (11) (2020) 1–19, http://dx.doi.org/10.2196/15293.

[51] R. Skaik, D. Inkpen, Suicide ideation estimators within canadian provinces using machine learning tools on social media text, Adv. Inf. Technol. 12 (4) (2021) 357–362, http://dx.doi.org/10.12720/jait.

[52] J. Du, Y. Zhang, J. Luo, Y. Jia, Q. Wei, C. Tao, H. Xu, Extracting psychiatric stressors for suicide from social media using deep learning, BMC Med. Inf. Decis. 18 (2) (2018) 77–87, http://dx.doi.org/10.1186/s12911-018-0632-8.

[53] T.H. Aldhyani, S.N. Alsubari, A.S. Alshebami, H. Alkahtani, Z.A. Ahmed, Detecting and analyzing suicidal ideation on social media using deep learning and machine learning models, Int. J. Env. Res. Public Heal. 19 (19) (2022) 1–16, http://dx.doi.org/10.3390/ijerph191912635.

[54] M. Gaur, V. Aribandi, A. Alambo, U. Kursuncu, K. Thirunarayan, J. Beich, J. Pathak, A. Sheth, Characterization of time-variant and time-invariant assessment of suicidality on reddit using c-ssrs, PloS One 16 (5) (2021) 1–21, http://dx.doi.org/10.1371/journal.pone.0250448.

[55] A. Malhotra, R. Jindal, Multimodal deep learning based framework for detecting depression and suicidal behaviour by affective analysis of social media posts, EAI Endorsed Trans. Pervasive Heal. Technol. 6 (21) (2020) http://dx.doi.org/10.4108/eai.13-7-2018.164259.

[56] M.M. Tadesse, H. Lin, B. Xu, L. Yang, Detection of suicide ideation in social media forums using deep learning, Algorithms 13 (1) (2019) 1–19, http://dx.doi.org/10.3390/a13010007.

[57] S. Renjith, A. Abraham, S.B. Jyothi, L. Chandran, J. Thomson, An ensemble deep learning technique for detecting suicidal ideation from posts in social media platforms, J. King Saud Univ. Comp. 34 (10) (2022) 9564–9575, http://dx.doi.org/10.1016/j.jksuci.2021.11.010.

[58] R. Sawhney, H. Joshi, S. Gandhi, D. Jin, R.R. Shah, Robust suicide risk assessment on social media via deep adversarial learning, J. Am. Med. Inf. Assoc. 28 (7) (2021) 1497–1506, http://dx.doi.org/10.1093/jamia/ocab031.

[59] T. Zhang, A.M. Schoene, S. Ananiadou, Automatic identification of suicide notes with a transformer-based deep learning model, Internet Interv. 25 (1) (2021) 1–8, http://dx.doi.org/10.1016/j.invent.2021.100422.

[60] A.G. Bermúdez, D. Carramiñana, A.M. Bernardos, L. Bergesio, J.A. Besada, Afusion architecture to deliver multipurpose mobile health services, Comput. Biol. Med. 173 (2024) 108344, http://dx.doi.org/10.1016/j.compbiomed.2024.108344.

[61] U. Naseem, M. Khushi, J. Kim, A.G. Dunn, Hybrid text representation for explainable suicide risk identification on social media, IEEE Trans. Comput. Soc. Syst. 11 (4) (2024) 4663–4672, http://dx.doi.org/10.1109/TCSS.2022.3184984.

[62] Y. Ma, Social media-based suicide risk detection via social interaction and posted content, in: International Conference on Artificial Intelligence and Information Systems, 2021, pp. 1–5, http://dx.doi.org/10.1145/3469213.3470345.

[63] L. Cao, H. Zhang, L. Feng, Building and using personal knowledge graph to improve suicidal ideation detection on social media, IEEE Trans. Multimed. 24 (1) (2020) 87–102, http://dx.doi.org/10.1109/TMM.2020.3046867.

[64] M. M, E. Cambria, Will affective computing emerge from foundation models and general artificial intelligence? a first evaluation of chatgpt, IEEE Intell. Syst. 38 (2) (2023) 15–23, http://dx.doi.org/10.1109/MIS.2023.3254179.

[65] E. Cambria, X. Zhang, R. Mao, M. Chen, K. Kwok, Senticnet 8: Fusing emotion ai and commonsense ai for interpretable, trustworthy, and explainable affective computing, in: International Conference on Information Reuse and Integration for Data Science, 2024, pp. 1–20.

[66] Q. Li, J. Li, J. Sheng, S. Cui, J. Wu, Y. Hei, H. Peng, S. Guo, L. Wang, A. Beheshti, P.S. Yu, Asurvey on deep learning event extraction: Approaches and applications, IEEE Trans. Neural Netw. Learn. Syst. 35 (5) (2024) 6301–6321, http://dx.doi.org/10.1109/TNNLS.2022.3213168.

[67] S. Tanwar, J. Singh, Resnext50 based convolution neural network-long short term memory model for plant disease classification, Multimedia Tools Appl. 82 (19) (2023) 29527–29545, http://dx.doi.org/10.1007/s11042-023-14851-x.

[68] H. Duan, Y. Long, S. Wang, H. Zhang, C.G. Willcocks, L. Shao, Dynamic unary convolution in transformers, IEEE Trans. Pattern Anal. Mach. Intell. 45 (11) (2023) 12747–12759, http://dx.doi.org/10.1109/TPAMI.2022.3233482.

[69] G. Anand, R. Nayak, Delta: deep local pattern representation for time-series clustering and classification using visual perception, Knowl.-Based Syst. 212 (2021) 1–16, http://dx.doi.org/10.1016/j.knosys.2020.106551.

[70] S. Gao, M. Alawad, M.T. Young, J. Gounley, N. Schaefferkoetter, H.J. Yoon, X.-C. Wu, E.B. Durbin, J. Doherty, A. Stroup, et al., Limitations of transformers on clinical text classification, IEEE J. Biomed. Heal. Inf. 25 (9) (2021) 3596–3607, http://dx.doi.org/10.1109/JBHI.2021.3062322.

[71] K.A. Sathi, M.K. Hosain, M.A. Hossain, A.Z. Kouzani, Attention-assisted hybrid 1d cnn-bilstm model for predicting electric field induced by transcranial magnetic stimulation coil, Sci. Rep. 13 (1) (2023) 1–12, http://dx.doi.org/10.1038/s41598-023-29695-6.

[72] C. Yan, J. Liu, W. Liu, X. Liu, Research on public opinion sentiment classification based on attention parallel dual-channel deep learning hybrid model, Eng. Appl. Artif. Intell. 116 (2022) 1–16, http://dx.doi.org/10.1016/j.engappai.2022.105448.

[73] L. Alzubaidi, J. Zhang, A.J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al-Shamma, J. Santamaría, M.A. Fadhel, M. Al-Amidie, L. Farhan, Review of deep learning: Concepts, cnn architectures, challenges, applications, future directions, J. Big Data 8 (2021) 1–74, http://dx.doi.org/10.1186/s40537-021-00444-8.

[74] H. Fan, X. Yu, Y. Yang, M. Kankanhalli, Deep hierarchical representation of point cloud videos via spatio-temporal decomposition, IEEE Trans. Pattern Anal. Mach. Intell. 44 (12) (2021) 9918–9930, http://dx.doi.org/10.1109/TPAMI.2021.3135117.

[75] K. Wang, Y. Wang, S. Zhang, Y. Tian, D. Li, Slms-ssd: Improving the balance of semantic and spatial information in object detection, Expert Syst. Appl. 206 (2022) 1–10, http://dx.doi.org/10.1016/j.eswa.2022.117682.

[76] M.B. McDermott, B. Yap, P. Szolovits, M. Zitnik, Structure-inducing pre-training, Nat. Mach. Intell. 5 (6) (2023) 612–621, http://dx.doi.org/10.1038/s42256-023-00647-z.

[77] N. Ding, Y. Qin, G. Yang, F. Wei, Z. Yang, Y. Su, S. Hu, Y. Chen, C.-M. Chan, W. Chen, et al., Parameter-efficient fine-tuning of large-scale pre-trained language models, Nat. Mach. Intell. 5 (3) (2023) 220–235, http://dx.doi.org/10.1038/s42256-023-00626-4.

[78] P. Schramowski, C. Turan, N. Andersen, C.A. Rothkopf, K. Kersting, Large pre-trained language models contain human-like biases of what is right and wrong to do, Nat. Mach. Intell. 4 (3) (2022) 258–268, http://dx.doi.org/10.1038/s42256-022-00458-8.

[79] Q. Zhang, J. Xiao, C. Tian, J. Chun-Wei Lin, S. Zhang, Arobust deformed convolutional neural network (cnn) for image denoising, CAAI Trans. Intell. Technol. 8 (2) (2023) 331–342, http://dx.doi.org/10.1049/cit2.12110.

[80] X. Wang, J. Liu, S. Chen, G. Wei, Effective light field de-occlusion network based on swin transformer, IEEE Trans. Circuits Syst. Video Technol. 33 (6) (2023) 2590–2599, http://dx.doi.org/10.1109/TCSVT.2022.3226227.

[81] Z. Wu, H. Rockwell, Y. Zhang, S. Tang, T.S. Lee, Complexity and diversity in sparse code priors improve receptive field characterization of macaque v1 neurons, PLoS Comput. Biol. 17 (10) (2021) 1–21, http://dx.doi.org/10.1371/journal.pcbi.1009528.

[82] N. Komati, Suicide and depression detection, 2021, https://www.kaggle.com/datasets/nikhileswarkomati/suicide-watch.

[83] A. Chadha, B. Kaushik, Ahybrid deep learning model using grid search and cross-validation for effective classification and prediction of suicidal ideation from social network data, New Gener. Comput. 40 (4) (2022) 889–914, http://dx.doi.org/10.1007/s00354-022-00191-1.