IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661, p-ISSN: 2278-8727, Volume 27, Issue 2, Ser. 1 (Mar. – Apr. 2025), PP 57-68
www.iosrjournals.org

A Literature Survey On Sentiment Analysis Using Image Processing

Srijan Sen [1], Shourov Sikder[1]
(Department Of Information Technology, Institute Of Engineering & Management, India)

Abstract:
Sentiment analysis, the automated process of determining emotions and opinions expressed in content, has evolved to encompass visual media, allowing for a deeper understanding of sentiments conveyed in images. This paper explores the application of image processing techniques, coupled with Python programming, to conduct sentiment analysis by extracting emotional cues from visual data.
The study begins with an overview of the significance and challenges of sentiment analysis in images, emphasizing the need for advanced tools to analyze the ever-growing volume of visual content on the internet. Leveraging Python's rich ecosystem of libraries, the paper delves into the technical aspects of sentiment analysis using image data.
Key components of this research include the utilization of Convolutional Neural Networks (CNNs) for feature extraction, pre-trained models for sentiment recognition, and the development of custom datasets to train and validate sentiment analysis models. Python libraries like TensorFlow and Keras provide a robust framework for building and deploying deep learning models.
The paper discusses the ethical considerations related to image-based sentiment analysis, addressing concerns about privacy, bias, and cultural nuances. It also explores the potential applications of this technology, ranging from brand sentiment analysis in marketing to monitoring public sentiment on social media platforms.
Furthermore, the study identifies challenges and opportunities in the field, paving the way for future research endeavors. By bridging the domains of computer vision, natural language processing, and machine learning, sentiment analysis by image processing in Python opens up new avenues for understanding the emotional impact of visual content in an increasingly digital and visually driven world.
Key Word: *Sentiment Analysis, Image Processing, Convolutional Neural Networks (CNNs), Facial Expression Recognition,Python Libraries (TensorFlow, Keras, OpenCV), Deep Learning, FER 2013.*

Date of Submission: 27-02-2025 Date of Acceptance: 07-03-2025

I. Introduction

Today's world is completely digital, and the usage of images and videos has grown a lot. Visual content is posted by people a lot on social media, internet forums, and messaging apps in order to communicate their thoughts, feelings, and experiences. So, with the rise in usage, it is now necessary to create sophisticated technology that can comprehend and analyze the emotions embedded within these images. Historically, sentiment analysis has been employed to analyze emotions and opinions from text, e.g., reviews, comments, or articles. This has assisted businesses, researchers, and developers in understanding public opinions and human emotions better. Nevertheless, with the advancement of digital communication, text alone is no longer sufficient. Most online communication occurs through images and videos, which also possess a lot of emotional and sentimental value. This requires an enhanced sentiment analysis technique that moves beyond text to include visual materials.

This project centers on uniting sentiment analysis with image processing, particularly examining facial expressions in order to deduce emotions. Facial expressions are an inherent and global method by which humans exhibit emotions like happiness, sadness, anger, surprise, and fear. In contrast to words, which can be misunderstood or lost in translation, facial expressions give a clear and strong indication of human emotions. Through the use of Python programming and some image processing methods, this project seeks to create a system that can effectively identify facial expressions and determine the corresponding emotions.

The capacity to read emotions from pictures has a lot of uses. Social networks can utilize it to read people's responses and interactions. Companies can use it to evaluate customer moods in advertising or online communication. Mental health tracking is also possible, where the technology can help experts analyze emotional state by looking at facial expressions. It can also improve human-computer interaction, where machines can better address human emotions naturally and intelligently.

With this study, we hope to fill the void in sentiment analysis by incorporating visual emotion recognition in current frameworks. By cracking the code of implicit emotional signals encoded in images, our project hopes to better comprehend how emotions are visually expressed and how they could be analyzed through contemporary technology.

II. Motivation:

The motivation behind this research stems from the recognition that images play a pivotal role in contemporary communication, offering a rich and nuanced medium for expressing emotions. While textual sentiment analysis has made substantial strides, the visual dimension of sentiments, particularly through facial expressions, remains largely untapped. Understanding the emotional content conveyed by an individual's face in images holds the key to unlocking deeper insights into user experiences, customer feedback, and social interactions.

III. Literature Review

Sentiment analysis has largely been text-oriented, but as the use of images on social media continues to grow, researchers are now turning their attention to how emotions can be analyzed from visual content. Research indicates that using text and images together increases accuracy, while deep learning methods such as CNNs and RNNs have been effective in image sentiment analysis. Facial expression recognition is also receiving interest as an accurate method of detecting emotions from pictures. This review points to major research in these fields, assisting in comprehending how visual sentiment analysis could be enhanced with sophisticated methods.

Wang & Li (2015) focuses on understanding human sentiments from a large-scale collection of internet images using both image features and contextual social network information. It extends the advances in text-based sentiment prediction to the challenge of predicting sentiments behind images. The proposed method significantly outperformed existing state-of-the-art methods in sentiment analysis on two large-scale datasets, achieving improvements of 10% and 6% in precision accuracy, respectively. The paper emphasizes the significance of unsupervised sentiment analysis, particularly given the challenges associated with obtaining labeled datasets for social media images. The proposed unsupervised approach allows for the utilization of vast amounts of unlabeled data, which is crucial for improving sentiment prediction in this domain. The method also showed that the performance on Instagram was worse than on Flickr, potentially due to the prevalence of "picture filters" affecting the discriminative ability of low-level visual features. The paper highlights that visual features such as color histograms and brightness lack the semantic depth required for effective sentiment prediction, making unsupervised sentiment analysis inherently more challenging than supervised method. [1]

The study conducted in [2] investigates the role of Twitter as a significant platform for sentiment analysis, emphasizing its ability to provide organizations with rapid insights into public opinions regarding their brands and services. It defines sentiment as an individual's emotional state, which can influence their actions and interactions with companies. The findings reveal that while human evaluators perceived the sentiment analysis as valuable, there were notable discrepancies between human and algorithmic evaluations, with algorithms demonstrating more consistent results. However, the study identifies several limitations that may impact the reliability of its findings. The sample size of observations was relatively small and did not meet the gold standards for sentiment classification, which could have skewed the results. Additionally, the human evaluators had limited experience and knowledge about the company, which may have led to misinterpretations of tweets, especially those not directly targeting the company but rather addressing external issues like government actions or competitors. The high variation in human classification is attributed to the subjective nature of their judgments and personal attributes, further complicating the reliability of sentiment analysis. The study suggests that future research should involve a larger dataset and more experienced evaluators to improve the accuracy of sentiment classification, thereby enhancing the overall understanding of how sentiments are expressed on social media platforms like Twitter. [2]

The paper [3] addresses the growing need for effective sentiment analysis in the digital age, where opinions are increasingly expressed through various modalities, including text, images, audio, and video. The study proposes a system that automatically recognizes facial expressions from images and classifies emotions, utilizing the Viola-Jones face detection algorithm for localization and employing feature extraction methods such as Zernike moments, LBP, and DCT transform. The results indicate that the system effectively combines different feature vectors using a subset feature selection algorithm, enhancing the performance of recognition and classification processes. The classifiers used in the study include SVM, Random Forest, and KNN, with experiments conducted on the JAFFE database to evaluate the system's performance. However, the paper acknowledges certain limitations, such as the reliance on a specific dataset, which may not encompass the diversity of facial expressions across different demographics and environments. Additionally, the performance may be affected by factors like lighting conditions and occlusions, which can hinder accurate facial feature detection. Overall, while the proposed system shows promise in improving sentiment analysis through facial expression recognition, further research is needed to address these limitations and enhance its applicability in real-world scenarios.

The paper [4] explores the burgeoning field of sentiment analysis, particularly focusing on images shared across social media platforms, where user's express emotions through visual content. It highlights the challenges of analyzing unlabelled images and presents deep learning techniques, such as Deep Neural Networks (DNN), Convolutional Neural Networks (CNN), Region-based CNN (R-CNN), and Fast R-CNN, as effective solutions for this task. The study summarizes various models and their applications, emphasizing that CNNs outperform other methods in terms of accuracy and efficiency, achieving nearly 20% higher accuracy compared to DNN and R-CNN. However, the paper also identifies limitations, such as the high computational cost and memory requirements associated with R-CNN, which processes multiple regions of an image, and the need for extensive labeled datasets for training deep learning models effectively. [4]

This paper [5] investigates the sentiment analysis of Twitter data, specifically focusing on tweets related to the 2016 US election, employing the Valence Aware Dictionary for sEntiment Reasoner (VADER) to classify sentiments expressed in these tweets. The study aims to address the limitations of previous research, which predominantly utilized binary classification methods, by proposing a multi-classification system that can categorize sentiments into multiple classes. The findings reveal that VADER effectively classifies tweets into various sentiment categories, demonstrating good accuracy in detecting ternary and multiple classes of sentiment, thus showcasing its potential for nuanced sentiment analysis in social media contexts. However, the study acknowledges certain limitations, particularly the reliance on VADER, which, while effective, may not capture all the complexities of sentiment expressed in tweets. Additionally, the focus on a specific event, the 2016 US election, may limit the generalizability of the results to other contexts or topics. Overall, the research contributes to the growing field of sentiment analysis by highlighting the importance of multi-classification approaches and the need for further exploration of sentiment analysis tools to enhance accuracy and depth in understanding public opinion expressed on social media platforms. [5]

The paper [6] explores the sentiment of users regarding the political Ayodhya issue through the analysis of tweets, leveraging machine learning algorithms to classify the polarity of these tweets. The authors highlight the challenges posed by the voluminous and diverse nature of social media data, which complicates manual sentiment analysis. They implemented a systematic approach that includes data collection from Twitter, preprocessing to clean the data by removing irrelevant elements like URLs and stop-words, and feature extraction using the TF-IDF method to represent the tweets as vector features. The results indicate that the machine learning model effectively classifies the sentiment of tweets, providing insights into public opinion on the political issue at hand. However, the paper does not explicitly mention any limitations, which could include potential biases in the training dataset, the challenges of accurately interpreting sentiments expressed in informal language, and the need for a more extensive dataset to improve the model's accuracy and generalizability. Future work could involve utilizing frameworks like Hadoop for better handling of large-scale political Twitter data, enhancing the analysis capabilities further.[6]

The paper [7] explores the integration of text and image data for sentiment analysis on Twitter, addressing the limitations of previous studies that primarily focused on single modalities. The authors propose a novel multimodal sentiment analysis model that combines textual and visual content, recognizing the growing prevalence of images, memes, and GIFs in social media communication. The model employs SentiBank and SentiStrength for image sentiment scoring, while text sentiment is assessed using a hybrid approach that combines lexicon-based and machine-learning techniques. The results demonstrate a high accuracy of 91.32% for the proposed model when evaluated on a random multimodal tweet dataset, outperforming separate models that analyze text or images independently. However, the study acknowledges several limitations, including the constraints of the Computer Vision API used for text recognition, which may struggle with informal language, code-mixing, and handwritten text. Additionally, the model does not account for other modalities such as animated GIFs and memes, which represent an open area for future research. Overall, this research contributes significantly to the field of sentiment analysis by providing a comprehensive framework that captures the nuances of multimodal communication on social media platforms. [7]

The paper [8] provides a comprehensive introduction to the emerging field of Visual Sentiment Analysis, which focuses on understanding how images evoke emotions in viewers. It highlights the historical context of sentiment analysis, primarily rooted in text analysis, and points out the limited exploration of visual content in this domain. The study reviews significant advancements in the field, particularly from 2010 to 2019, and discusses various methodologies, emotional models, dataset definitions, and feature designs that have been developed to analyze sentiments in images. The results indicate that while there has been notable progress, challenges remain, particularly in the formalization of sentiment analysis for images, which includes factors like the sentiment holder and temporal aspects that are often overlooked. The paper also identifies limitations in current methodologies, such as the need for better filtering of textual information associated with images to avoid noise and inaccuracies in sentiment inference. Furthermore, it suggests future research directions, including the exploration of visual sentiment analysis in videos, which introduces additional complexities due to the temporal dimension and multimodal data. Overall, the paper serves as a reference for researchers interested in this field, providing insights into existing challenges and opportunities for further exploration. [8]

The paper [9] focuses on the growing importance of social media platforms, particularly Twitter, as a source for understanding public sentiment. It highlights how users express their opinions, emotions, and experiences through tweets, which can be valuable for organizations seeking customer feedback and satisfaction. The study employs Natural Language Processing (NLP) techniques, specifically the Natural Language Toolkit (NLTK) and VADER (Valence Aware Dictionary and Sentiment Reasoner), to analyze sentiments from tweets, categorizing them into positive, negative, and neutral sentiments. The methodology involves extracting a dataset of 3,000 tweets related to US airlines, which are initially in an unstructured format. The data undergoes a cleaning process to remove irrelevant elements such as URLs, usernames, and punctuation, ensuring that only the text is analyzed. The results indicate that out of the analyzed tweets, there were 6,313 positive, 5,243 negative, and 2,930 neutral tweets, showcasing a diverse range of sentiments expressed by users. However, the study acknowledges limitations, such as the reliance on Twitter as the sole data source, which may not fully represent broader public sentiment. Additionally, the inherent challenges of analyzing unstructured data and the potential biases in user-generated content could affect the accuracy of sentiment classification. Overall, the paper emphasizes the significance of sentiment analysis in understanding customer perspectives and guiding organizational strategies. [9]

The paper titled [10] focuses on developing a Deep Convolutional Neural Network (DCNN) model aimed at classifying five distinct human facial emotions. The introduction highlights the significance of emotion recognition systems, which leverage advancements in artificial intelligence and deep learning to meet real-time human needs, particularly in applications like feedback analysis and face unlocking. The study employs a manually collected image dataset for training, testing, and validation of the model, which is optimized using the Adam algorithm and categorical cross-entropy as the loss function. The results indicate that the model achieves a training accuracy and validation accuracy of 78.04%, suggesting a good fit and generalization to the data. Additionally, the model's performance is compared with existing models, revealing that it outperforms some, while also indicating potential for future enhancements, such as analyzing emotional changes through video sequences for real-time applications. However, limitations include the model's reliance on a specific dataset, which may affect its generalizability to diverse real-world scenarios, and the need for further testing against a broader range of emotional expressions and conditions to validate its robustness. Overall, the research contributes valuable insights into the capabilities of DCNNs in emotion recognition, while also identifying areas for future exploration and improvement. [10]

In the paper [11], the authors introduce a novel framework aimed at enhancing image sentiment analysis by focusing on mid-level attributes. The motivation behind this work stems from the increasing prevalence of images on social media platforms like Twitter, where images constitute a significant portion of shared content. The authors highlight that while textual sentiment analysis has been extensively explored, visual sentiment analysis remains underdeveloped. They propose that images carry sentiment that can be interpreted through mid-level attributes, which offer a more interpretable approach compared to traditional low-level feature analysis. The Sentribute framework is structured into several sections: it begins with an overview of the proposed method, followed by detailed discussions on low-level feature extraction, mid-level attribute generation, and the integration of facial sentiment recognition to improve prediction accuracy. The paper concludes with a summary of findings and potential future research directions, emphasizing the need for further exploration in this area. The results of the Sentribute framework demonstrate its effectiveness in predicting image sentiment. The authors conducted empirical studies using a dataset of 810 images sourced from Twitter, comparing their method against state-of-the-art techniques that primarily rely on low-level features and textual information. The findings indicate that the Sentribute framework, which incorporates mid-level attributes and eigenface-based emotion detection, significantly enhances prediction accuracy, particularly for images containing faces. The use of an asymmetric bagging approach also addresses the challenges posed by unbalanced datasets, leading to improved classification performance. The results show that the precision of sentiment classification using both Linear SVM and Logistic Regression algorithms outperforms recall, indicating a smaller number of false positives and a higher detection rate for true positives. Overall, the study highlights the potential of mid-level attributes in visual sentiment analysis and suggests that combining these attributes with textual sentiment analysis could yield even more robust results. [11]

The paper [12], addresses the evolution of sentiment analysis in the context of social media, where users increasingly share opinions through image-text posts, moving beyond traditional text-only formats. This shift necessitates a more complex approach to sentiment analysis, leading to the development of a multimodal sentiment analysis method that effectively integrates visual and textual content. The proposed approach utilizes mid-level visual features from the SentiBank method, alongside textual and social features, to enhance sentiment analysis by exploring the correlation between images and text. The results of extensive experiments demonstrate the superior performance of this new method compared to conventional approaches, indicating its effectiveness in capturing the nuances of sentiment in image-text posts. However, the study acknowledges limitations, such as the potential for misleading sentiment analysis due to the inconsistency between images and text in user-generated content, which can lead to challenges in accurately interpreting sentiment. Additionally, the reliance on low-level visual features may still result in a semantic gap, suggesting that further refinement in feature extraction could enhance the model's performance**.** [12]

This study [13] focuses on the classification of human facial expressions in real-time images, utilizing deep learning methods and image processing techniques. The research highlights the potential applications of such technology, including mood analysis in group photos and context-aware image access, where specific emotions can be filtered from a database. The study classifies seven different emotions: happiness, sadness, surprise, disgust, anger, fear, and neutral, using various machine learning and deep learning architectures, including k-Nearest Neighbors, Support Vector Machines, AlexNet, ResNet, DenseNet, and Inception, applied to datasets like FER2013, JAFFE, and CK+. The results indicate that the highest accuracy was achieved in recognizing neutral expressions, while the lowest accuracy was noted for surprise and disgust expressions. The study also found that real-time applications performed better when using the FER2013 dataset with the ResNet architecture, particularly when employing Histograms of Oriented Gradients (HOG) for feature extraction. However, the study faced limitations, such as challenges in real-time recognition for images above 720p resolution, a restricted detection range of 50 to 100 cm from the webcam, and performance being affected by lighting conditions and camera angles. These factors highlight the need for optimal environmental conditions to enhance the system's performance.[13]

The paper [14] addresses the growing need for sentiment analysis in visual content, particularly in social media, where images and videos are increasingly used to express opinions. The authors propose a Convolutional Neural Network (CNN) architecture specifically designed for visual sentiment analysis, leveraging a large dataset of half a million images labeled through a baseline sentiment algorithm. They employ a progressive training strategy to refine the model using noisy data and enhance performance on Twitter images through domain transfer with a smaller set of manually labeled images. The results demonstrate that their CNN outperforms existing algorithms in image sentiment analysis, showcasing its effectiveness in handling weakly labeled data and improving generalizability across domains. However, the study acknowledges limitations, such as the inherent challenges of image sentiment analysis compared to object recognition, due to the subjective nature of sentiment and the need for extensive labeled datasets, which can be labor-intensive to compile. Additionally, while the proposed methods show promise, the reliance on machine-generated labels may introduce noise that could affect the model's accuracy. [14]

The paper [15] introduces a novel approach to visual sentiment analysis, focusing on the emotional responses elicited by images and videos. It highlights the limitations of existing models that primarily utilize global visual features, arguing that local image regions are crucial for understanding human emotional responses. The proposed model employs an attention mechanism to automatically discover relevant local regions within images, which enhances the sentiment classification process. Experimental results demonstrate that this model not only identifies sentimental local regions effectively but also outperforms current state-of-the-art algorithms in visual sentiment analysis. However, the study acknowledges certain limitations, such as the reliance on the performance of the visual attribute detector, which is not the main focus of the research. The effectiveness of the proposed model could be further improved by refining the visual attributes used in the analysis. Overall, while the research presents a significant advancement in the field, it also opens avenues for future work, particularly in enhancing the robustness of visual attribute detection and exploring larger datasets for better sentiment classification. [15]

The paper [16] introduces the field of Image Sentiment Analysis, highlighting its significance in understanding emotions conveyed through images, which has been less explored compared to text-based sentiment analysis. It provides a comprehensive summary of current research, discussing various emotional representation models, datasets, and features used in the analysis. The results indicate that while significant progress has been made, challenges remain, particularly in achieving consensus on emotional classification schemes and the lack of universal benchmark datasets, which complicates result comparisons. Additionally, the paper notes that many existing studies face limitations such as arbitrary emotional outputs and the need for large-scale datasets to train effective machine learning models. Future work is suggested to address these challenges by exploring new methods, features, and datasets, ultimately aiming to enhance the accuracy and applicability of image sentiment analysis systems. [16]

The paper [17] addresses the challenge of sentiment analysis using a large dataset of user-generated content from social media, specifically focusing on images and their associated textual descriptions. The authors highlight the increasing importance of visual sentiment analysis, especially as users often share images with little or no text, making traditional text-based sentiment analysis insufficient. They propose a novel approach that leverages the sentiment polarity of textual content to train a visual sentiment classifier, utilizing over 3 million tweets containing both text and images. The results demonstrate that despite the noisy and weak correlation of text with image content, the proposed deep Convolutional Neural Network effectively predicts sentiment polarity for unseen images, marking a significant advancement in the field. However, the study acknowledges limitations, such as the inherent subjectivity of sentiment interpretation and the challenges posed by the affective gap between image features and their emotional content, which complicates the visual sentiment analysis process. Overall, this work contributes to the understanding of sentiment in uncontrolled environments, paving the way for future research in multimedia sentiment analysis. [17]

The paper [18] aims to create subjective descriptions of images, addressing the challenge of developing a model that can understand both image content and its natural language representation without relying on predefined templates or categories. The authors propose a Kernel Ridge Regression model to map images to text, focusing on three sentiment categories: positive, neutral, and negative. They utilize two types of image features: simple RGB pixel values and more complex features derived from deep learning techniques. The experimental evaluation is conducted on a Twitter dataset that includes both images and associated text, revealing that the proposed mapping performs better for positive sentiments compared to neutral and negative ones. Additionally, the results indicate that deep learning features outperform RGB pixel values across all sentiment categories, especially with larger training sets. However, the study acknowledges limitations, such as the varying performance across different sentiment categories and the potential constraints of the model in capturing the full richness of human emotional responses to images. Overall, the research contributes to the field of sentiment analysis by advancing the understanding of image-text relationships and suggesting future directions for enhancing model capabilities. [18]

The paper [19] introduces an Attention-based Heterogeneous Relational Model (AHRM) aimed at enhancing multi-modal sentiment analysis for social images, which often combine visual and textual data along with social attributes like tags and user interactions. AHRM employs a progressive dual attention mechanism to effectively capture the emotional correlations between images and their accompanying text, while also integrating social context through a heterogeneous relation network that extends Graph Convolutional Networks (GCN). The results demonstrate that AHRM outperforms traditional unimodal methods and other baseline models, achieving superior sentiment classification accuracy by leveraging both content and structural information. However, the study acknowledges limitations, such as the potential for performance gaps when relying solely on visual or textual content, indicating that neither modality alone suffices for accurate sentiment inference. Additionally, while AHRM shows promise, the reliance on social context may not be universally applicable across all datasets, suggesting that further research is needed to explore its effectiveness in diverse scenarios. **[19]**

The paper evaluates various deep learning techniques for sentiment analysis specifically using Twitter data, highlighting the growing importance of these methods in addressing complex problems in natural language processing (NLP). It compares convolutional neural networks (CNN) and long short-term memory (LSTM) networks, alongside different word embedding systems like Word2Vec and GloVe, to assess their performance under a unified testing framework. The results indicate that using multiple CNN configurations combined with LSTM networks significantly enhances performance, achieving improvements of 3%-6% across various setups. Additionally, the GloVe embedding system consistently outperformed Word2Vec, likely due to its training on a larger dataset, which contributed to better vectorization of words. However, the study also notes limitations, such as the minimal performance gain from separating text input into regions and the lack of advantage from using bidirectional LSTM networks over standard LSTM configurations, which may be attributed to the nature of the data. Overall, the paper emphasizes the critical role of dataset quality and network configuration in achieving optimal results in sentiment analysis.[20]

IV. Methodology

The methodology presented in [1] focuses on a dual approach to sentiment analysis by integrating both visual and textual information. Initially, the authors extract mid-level visual features from images using large-scale visual attribute detectors, resulting in a feature dimension of 1200. Concurrently, textual features are derived from user profiles, image captions, and comments, which undergo preprocessing steps such as stop word removal and stemming to enhance their quality. The sentiment prediction problem is formulated in two scenarios: supervised and unsupervised, allowing for flexibility in model training. An optimization algorithm is developed to iteratively update the model components, including visual and textual sentiment label spaces, until convergence is achieved. The authors employ an alternating multiplicative updating scheme to optimize the objective function, ensuring that both visual and textual contributions are effectively integrated into the sentiment analysis framework. This comprehensive methodology allows for a more accurate sentiment labeling by leveraging the strengths of both modalities, addressing the limitations of relying solely on either visual or textual data

The methodology of the study involved a comprehensive approach to collecting and analyzing Twitter data from a Finnish software company. A total of 509 tweets were gathered using the Nemo Sentiment and Data Analyzer tool, which is specifically designed for sentiment analysis in the Finnish language. To assess the reliability of the sentiment classifications, the researchers calculated Krippendorff's alpha, a statistical measure that evaluates the agreement among different evaluators. This was essential for understanding the consistency of sentiment classifications across human evaluators and automated tools. Additionally, the perceived value of the sentiment analyses was explored through a workshop with company representatives, including a business unit manager, a key account manager, and a marketing specialist, who provided feedback on the analysis results and their relevance for business decision-making. The study compared the classifications made by human evaluators against those generated by the automated tools, specifically focusing on the performance of SentiStrength and Nemo Sentiment and Data Analyzer, which utilize different algorithms for sentiment classification. [2]

The methodology for the facial expression recognition and emotion classification system in paper [3] is structured into several key phases that collectively enhance the process of sentiment analysis. Initially, the system utilizes the Viola-Jones face detection algorithm, which is crucial for accurately identifying and localizing faces in input images, thereby setting the foundation for effective analysis. Once the faces are detected, the next phase involves feature extraction, where three distinct techniques are employed: Zernike moments, Local Binary Patterns (LBP), and Discrete Cosine Transform (DCT). These methods are designed to capture significant facial features that correlate with various emotional states. Following feature extraction, the methodology incorporates a subset feature selection algorithm to combine the different feature vectors, optimizing the feature set to enhance the performance of the recognition and classification processes. The final phase of the methodology involves classification, where the system applies multiple classifiers, including Support Vector Machine (SVM), Random Forest, and K-Nearest Neighbors (KNN), to categorize the recognized facial expressions into specific emotional classes. This comprehensive approach not only addresses the challenges posed by varying facial expressions but also ensures accurate emotion classification. The effectiveness of the proposed methodology is validated through experiments conducted on the JAFFE database, demonstrating its capability to automate facial expression recognition and improve sentiment analysis. Overall, the methodology integrates advanced techniques in face detection, feature extraction, and classification, creating a robust system for emotion analysis. [3]

The methodology employed in the paper [4] revolves around leveraging various deep learning architectures to analyze sentiments conveyed through images. The authors begin by addressing the challenge of predicting sentiments from unlabelled images, which is a significant hurdle in the field. To tackle this, they explore several deep learning models, including Deep Neural Networks (DNN), Convolutional Neural Networks (CNN), Region-based CNN (R-CNN), and Fast R-CNN, each with its unique approach to feature extraction and classification. The paper emphasizes the effectiveness of CNNs, which are particularly adept at learning spatial hierarchies in images, thus enabling them to capture intricate patterns that correlate with emotional content. The authors also discuss the integration of techniques like SentiBank and SentiStrength to enhance sentiment detection by combining visual features with sentiment lexicons. Furthermore, the methodology includes empirical evaluations on various datasets, such as those sourced from social media platforms like Flickr, to validate the performance of the proposed models. The results indicate that CNNs significantly outperform DNNs and R-CNNs in terms of accuracy and processing efficiency, highlighting their suitability for image sentiment analysis. However, the paper also notes the limitations of R-CNN, particularly its high computational demands due to the processing of multiple image regions, which can hinder real-time applications. [4]

The methodology employed in [5] involves a systematic approach to sentiment analysis of Twitter data, specifically focusing on tweets related to the 2016 US presidential election. The process begins with data acquisition, where a total of 2,430 political tweets were collected using the Network Overview Discovery and Exploration for Excel (NodeXL) tool, utilizing relevant hashtags and keywords associated with the election candidates, Hillary Clinton and Donald Trump. Following data collection, a series of preprocessing steps were conducted using Python's Natural Language Toolkit (NLTK) to clean the data, which included converting tweets to lowercase, removing stop words, tokenizing the text, and stemming the words to prepare the dataset for analysis. The VADER Sentiment Analyzer was then applied to classify the preprocessed tweets into sentiment categories, utilizing a scoring rule to determine the overall sentiment polarity. This scoring rule allowed tweets to be classified into five sentiment classes: high positive, positive, neutral, negative, and high negative, based on their compound values. The study emphasizes the importance of using appropriate threshold values to accurately categorize the sentiments expressed in the tweets, ultimately demonstrating the effectiveness of VADER in multi-class sentiment analysis. [5]

The methodology employed in [6] involves several systematic steps. Initially, the authors collected tweets related to the political Ayodhya issue using the Twitter API, ensuring that the data gathered was relevant to the political discourse. The collected data underwent a rigorous preprocessing phase, where irrelevant elements such as URLs, user names, and stop-words were removed to enhance the quality of the dataset. This preprocessing step is crucial as it transforms raw data into a format that can be effectively utilized by classifiers. Following this, the authors utilized the TF-IDF (Term Frequency - Inverse Document Frequency) method for feature extraction, which allowed them to represent the tweets as vector features, assigning appropriate weights to different words based on their significance in the context of sentiment. The sentiment classification was performed using a supervised machine learning approach, where the model was trained on a newly created dataset using the Vader

Sentiment Analyzer. This comprehensive methodology not only facilitated the effective classification of tweet sentiments but also aimed to address the challenges posed by the diverse and informal nature of social media language. [6]

The methodology for [7] involves several key steps aimed at effectively analyzing both textual and visual content from Twitter. Initially, the process begins with data pre-processing, where the textual data is cleaned and transformed to facilitate relevant feature extraction. This includes decoding HTML entities, removing URLs, and eliminating retweet indicators, contractions, and unnecessary punctuation, ensuring that the text is standardized and devoid of noise. Following this, the paper proposes a hybrid technique that combines machine learning and lexicon-based approaches for context-aware textual sentiment analysis, allowing for a nuanced understanding of sentiment in tweets. For the visual content, sentiment scoring is performed using SentiBank and SentiStrength, which are specifically designed to evaluate the emotional content of images. The model then employs a convolutional neural network (CNN) to analyze the images, extracting features that contribute to sentiment determination. The final step involves aggregating the sentiment scores derived from both the textual and visual analyses, which is achieved by separating the text from the image using an optical character recognizer (OCR). This comprehensive approach allows the model to achieve a high-performance accuracy of 91.32% on the multimodal tweet dataset, demonstrating that the integration of both textual and visual features significantly enhances sentiment analysis compared to models that rely solely on one modality. Overall, the methodology emphasizes the importance of understanding the interplay between text and images in social media to accurately gauge sentiment.

The methodology outlined in the paper [8] emphasizes a structured approach to understanding how images evoke emotions. It begins with a review of existing literature, focusing on significant works published between 2010 and 2019, to identify the evolution of techniques and methodologies in the field. The paper categorizes the methodologies into three main areas: emotional models, dataset definitions, and feature designs, providing a critical viewpoint on the employed features and datasets. It discusses the integration of textual and visual information through multimodal embedding systems, which combine features from different modalities to create a common vector space, maximizing correlations between them. Additionally, the paper highlights the use of deep learning models trained for various visual inference tasks, which extract descriptive text related to the visual content. The methodology also addresses the importance of context in sentiment analysis, suggesting that knowledge of the context can significantly influence the interpretation of images. Overall, the paper presents a comprehensive framework for developing Visual Sentiment Analysis systems, while also identifying challenges such as the need for better dataset definitions and the subjective nature of textual descriptions associated with images. [8]

The methodology employed in the paper [9] revolves around Natural Language Processing (NLP) techniques to analyze sentiments from tweets. The process begins with the extraction of a large dataset of tweets, which are often unstructured, from Twitter using the Tweepy API. To facilitate sentiment analysis, the data is cleaned and processed to convert it into a structured format, removing irrelevant elements such as URLs, usernames, and punctuation. The study focuses on analyzing both individual tweets and larger datasets, allowing users to input text for sentiment analysis and receive categorized results of positive, negative, and neutral sentiments. The analysis utilizes the VADER sentiment analysis tool, which is a rule-based and lexicon-based approach, to determine the sentiment of each tweet. The results are then represented visually through various charts, such as pie charts and bar charts, to provide a clear understanding of the sentiments expressed. Overall, the methodology combines data extraction, processing, and sentiment classification to effectively analyze public opinions on social media platforms. [9]

The methodology employed in [10] involves the construction and training of a two-layer convolutional neural network (CNN) designed to classify five different facial emotions. The model architecture includes an input layer, followed by convolutional layers that utilize the ReLU (Rectified Linear Unit) activation function to process the feature maps generated from the input images, which are resized to 32 x 32 pixels. To mitigate overfitting, dropout layers are incorporated after each convolutional layer, and a pooling layer with a pool size of 2 x 2 is used to reduce the dimensionality of the feature maps without losing critical information. The model is built using the Keras deep learning library, which operates on top of TensorFlow, and is implemented in Python within a Jupyter Notebook environment. For model evaluation, the confusion matrix is generated using the Scikit-learn package, allowing for the assessment of accuracy, precision, sensitivity, specificity, and recall. The Adam optimizer is utilized to update the network weights, employing individual learning rates for each weight, while categorical cross-entropy serves as the loss function during training. This comprehensive approach ensures that the model is effectively trained and validated against the manually collected image dataset, ultimately achieving a training accuracy of 78.04% and demonstrating its potential for real-time emotion recognition applications. [10]

The Sentribute framework for image sentiment analysis employs a systematic approach centered on mid-level attributes. The methodology begins with the extraction of low-level features from images, which are then transformed into 102 mid-level attributes. This transformation is crucial as it enhances the interpretability of sentiment classification compared to relying solely on low-level features. A significant aspect of the methodology is the incorporation of eigenfaces for facial sentiment recognition. This technique is particularly effective in identifying extreme sentiments, allowing the framework to refine its predictions by combining results from both mid-level attributes and facial expressions. This dual approach aims to minimize errors in sentiment classification, thereby improving accuracy. The framework is empirically tested on a dataset of 810 images sourced from Twitter, with results benchmarked against existing methods that primarily utilize low-level features and textual data. To address challenges posed by unbalanced datasets, an asymmetric bagging technique is employed, further enhancing the robustness of the sentiment analysis process. Overall, the Sentribute framework represents a significant advancement in visual sentiment analysis by effectively integrating mid-level attributes and facial expression detection for improved prediction outcomes. [11]

In this study, a comprehensive approach was taken to develop a real-time facial expression recognition system using deep learning techniques. The research began with the selection of three widely recognized public datasets: FER2013, JAFFE, and CK+. These datasets were chosen for their relevance and extensive use in the field of facial expression recognition. The study employed various machine learning algorithms, including classical methods like k-Nearest Neighbors (k-NN) and Support Vector Machines (SVM), alongside advanced deep learning architectures such as AlexNet, ResNet, DenseNet, and Inception. The implementation was carried out using the Python programming language, which facilitated the integration of these algorithms into a cohesive software application. To evaluate the effectiveness of the proposed system, two distinct testing practices were conducted. The first practice involved testing all models with a set of 70 images that were not included in the training phase. This approach ensured an unbiased assessment of the models' performance. The second practice utilized images captured in real-time from a Logitech C270 webcam, allowing for a practical evaluation of the system's capabilities in a live environment. The feature extraction process was critical to the system's performance, employing the Histograms of Oriented Gradients (HOG) method to extract relevant features from the facial images. Following feature extraction, the study compared the classification performance of different algorithms, focusing on their accuracy rates across the datasets. Notably, the ResNet50 architecture demonstrated a high accuracy rate of 94.51% on the FER2013 dataset, indicating its effectiveness in recognizing facial expressions. The research also highlighted the importance of computational efficiency, as the system was designed to operate with minimal resources. The software was executed on a laptop equipped with an 8-core Intel Core i7 processor, utilizing only CPU cores for processing, which underscored the goal of achieving high performance without relying on more demanding GPU resources. Overall, the study aimed to demonstrate that real-time facial expression recognition can be effectively implemented using deep learning techniques, achieving high accuracy while maintaining operational efficiency. [12]

The methodology employed in this study involves a combination of image processing techniques and deep learning architectures to classify human facial expressions in real-time. The research utilizes a Logitech C270 720p webcam to capture images and videos, ensuring that the setup is in a well-lit environment to enhance accuracy. The study implements various machine learning methods, including k-Nearest Neighbors and Support Vector Machines, alongside deep learning architectures such as AlexNet, ResNet, DenseNet, and Inception, applied to datasets like FER2013, JAFFE, and CK+. For feature extraction, the Histograms of Oriented Gradients (HOG) method is employed, which is effective in capturing the essential features of facial expressions. The ResNet architecture, consisting of 50 layers, is particularly highlighted for its ability to achieve high accuracy rates, with ResNet50 demonstrating a 94.51% accuracy on the FER2013 dataset. The study also emphasizes the importance of using a window size of 128 x 128 pixels for face images, which is optimal for covering human faces during analysis **[3]**. The experiments were conducted using Python in a Spyder and Anaconda development environment on an 8-core Intel Core i7 processor, ensuring that the system could operate with minimal computational resources. Overall, the methodology integrates various techniques to create a robust system for real-time facial expression recognition, demonstrating the effectiveness of deep learning in this domain. [13]

The methodology employed in the paper revolves around the use of Convolutional Neural Networks (CNN) to effectively analyze visual sentiment. Initially, the authors designed a suitable CNN architecture tailored for image sentiment analysis, which consists of multiple convolutional layers followed by fully connected layers to predict sentiment labels. To address the challenge of large-scale weakly labeled training data, they collected a dataset of half a million images from Flickr, which were labeled using a baseline sentiment algorithm. Recognizing the noise inherent in machine-generated labels, the authors implemented a progressive training strategy that fine-tunes the CNN by retaining training instances with significant differences in predicted sentiment scores, thereby filtering out less informative data. Additionally, to enhance the model's performance on Twitter images, they introduced a domain transfer approach, utilizing a small set of manually labeled Twitter images to adapt the model to this specific domain. Extensive experiments were conducted on the manually labeled Twitter dataset to evaluate the effectiveness of the proposed methodology. The results indicated that the CNN architecture, combined with the progressive training and domain transfer strategies, significantly improved the performance of image sentiment analysis compared to competing algorithms, demonstrating the robustness and adaptability of the proposed approach in handling visual sentiment data. [14]

The methodology presented in the paper focuses on enhancing visual sentiment analysis by leveraging local image regions through an attention mechanism. Initially, the study employs pre-trained GloVe vectors to represent a set of 269 adjectives, which are crucial for sentiment classification. These vectors, being 300-dimensional, provide a robust semantic representation that aids in understanding the emotional context of the images. The model utilizes a logistic classifier built on top of these GloVe features, achieving an impressive accuracy of 95% through 5-fold cross-validation, indicating the effectiveness of the textual features in sentiment analysis. To further improve the model, the attention mechanism is introduced, which assigns relevance scores to different image regions based on their connection to the descriptive words. This approach allows the model to automatically identify and focus on the most sentimentally relevant local regions within an image, thereby enhancing the overall sentiment classification process. The attention model employs a bilinear function to evaluate the relevance scores, facilitating a more nuanced understanding of the relationship between the visual content and the associated textual descriptors. Overall, the methodology emphasizes the importance of local image features in visual sentiment analysis, marking a significant departure from traditional models that primarily rely on global image representations. [15]

The research in Image Sentiment Analysis employs a variety of techniques to classify and interpret the emotional content of images. It begins with the identification of emotional classes, which are often based on psychological theories, leading to the development of a Visual Sentiment Ontology (VSO) that categorizes images into specific emotional states such as anger, joy, and sadness. The analysis utilizes visual features extracted from images, which can include color histograms and object detection, to correlate these features with sentiment scores. For instance, a Convolutional Neural Network (CNN) is often fine-tuned to classify images into predefined emotional categories, leveraging pre-trained models to enhance performance. Additionally, the integration of metadata associated with images is explored to assign sentiment scores, allowing for a more nuanced understanding of the emotional context. The research also highlights the importance of datasets, noting that the lack of large-scale, standardized datasets poses a significant challenge for training effective models. Furthermore, the paper discusses the potential of using emoji and emoticons as supplementary tools for sentiment analysis, suggesting that these visual cues can enhance the understanding of emotional content in images. Overall, the methodologies employed in this field aim to bridge the gap between visual content and emotional interpretation, paving the way for more sophisticated sentiment analysis systems. [16]

The methodology employed in the study involves a cross-media learning approach to sentiment analysis, focusing on the integration of textual and visual data. Initially, the researchers collected a substantial dataset of over 3 million tweets containing both text and images using a streaming crawler from Twitter, which allowed them to access a random 1% sample of the global tweet stream over a six-month period from July to December 2016. The core of the methodology is the use of a bidirectional Long Short-Term Memory (bi-LSTM) network to generate document embeddings from the textual data, capturing long-range dependencies and contextual information. These embeddings serve as features for a Support Vector Machine (SVM) classifier, which is trained to classify the sentiment of images into three categories: positive, neutral, or negative. The researchers also utilized deep Convolutional Neural Networks (CNNs) for visual classification, leveraging their ability to learn hierarchical representations from image data. The training process involved using the sentiment polarity derived from the textual content to label the corresponding images, thus creating a training set for the visual classifier. The methodology emphasizes the importance of combining textual sentiment analysis with visual data to enhance the accuracy of sentiment predictions, despite the challenges posed by the noisy nature of social media content and the subjective interpretation of sentiments. Overall, this innovative approach demonstrates the potential of cross-media learning in effectively addressing the complexities of sentiment analysis in uncontrolled environments. [17]

The methodology employed in this study revolves around the use of a Kernel Ridge Regression model to facilitate the mapping of images to text, specifically in the context of sentiment analysis. The authors begin by selecting two distinct types of image features for their analysis: the first being simple RGB pixel-values, which represent the raw color information of the images, and the second being more sophisticated features extracted through deep learning techniques, which are capable of capturing complex patterns and representations within the images. To construct the text features, a bag-of-words model is utilized, allowing the researchers to represent the textual data in a format suitable for analysis. The study focuses on three sentiment categories—positive, neutral, and negative—enabling a nuanced understanding of how images can convey different emotional tones. The experimental evaluation is conducted on a real-world dataset sourced from Twitter, which includes both images and their corresponding text, along with the sentiment associated with each sample. The results of the experiments reveal that the proposed mapping technique performs particularly well for positive sentiment, outperforming the neutral and negative categories. Furthermore, the findings indicate that the deep learning-derived features consistently yield better performance across all sentiment categories compared to the simpler RGB pixel-value features, especially when larger training sets are employed. This comprehensive approach not only advances the understanding of image-to-text mapping but also highlights the potential of combining traditional regression techniques with modern deep learning methods in sentiment analysis. [18]

The proposed framework begins with Single-modal Representation Learning, where distinct representations for images and text are derived separately, allowing for a focused understanding of each modality's features. Following this, the Progressive Dual Image-Text Attention mechanism is introduced, which captures the intricate correlations between images and their corresponding textual descriptions. This is achieved through two innovative attention strategies: channel attention, which identifies significant semantic patterns in the image channels, and region attention, which highlights emotionally relevant areas within the images. By integrating these attentions, the model effectively combines visual and textual features into a cohesive joint representation. The next step involves Heterogeneous Relation Fusion, where social relations are utilized to construct a heterogeneous relation network. This network extends the capabilities of Graph Convolutional Networks (GCN) to aggregate contextual information from social interactions, enhancing the quality of the learned image representations. Finally, the sentiment prediction is conducted using the enriched joint representations, leading to improved classification outcomes. The entire process emphasizes the importance of both content and social context in achieving accurate sentiment analysis, showcasing the model's ability to leverage multi-modal data effectively. [19]

In this study, the authors present a comprehensive approach to sentiment analysis using Twitter data, focusing on various deep learning techniques. The dataset utilized for this research is derived from the international workshop on semantic evaluation (SemEval), which is known for its extensive collection of Twitter data. The authors explore different word embedding models, specifically Word2Vec and GloVe, to create word vectors that serve as input for the neural networks. The Word2Vec model is configured using the Continuous Bag of Words (CBOW) approach, generating 25-dimensional vectors while discarding words that appear less than five times in the dataset. GloVe, on the other hand, employs pretrained vectors created from a significantly larger corpus of 2 billion tweets, resulting in 25-dimensional vectors as well. The study also discusses the creation of sentence vectors by concatenating word vectors from tweets, ensuring that each tweet is standardized to a length of 40 words through repetition or truncation as necessary.The authors implement various deep neural network configurations, primarily focusing on convolutional neural networks (CNN) and long short-term memory (LSTM) networks. They propose two distinct CNN configurations that utilize different word embeddings, which are subsequently combined in a random forest classifier for enhanced performance. Notably, the study excludes GRU networks and RCNNs from the analysis, as they yield similar results to LSTM networks and CNNs. The evaluation process involves rigorous testing of these configurations to compare their effectiveness in sentiment analysis, ultimately aiming to clarify the advantages and limitations of each method within a single framework. [20]

V. Discussion

The following table shows the nomenclature of short forms:

AI	Artificial Intelligence
CNN	Convolutional Neural Networks
RNN	Recurrent Neural Network
LBP	Local Binary Pattern
DCT	Discrete Cosine Transform
SVM	Support Vector Machine
KNN	K-Nearest Neighbors
FER2013 Dataset	Facial Emotion Recognition Dataset
JAFFE Dataset	Japanese Female Facial Expression Dataset
DNN	Deep Neural Networks
R-CNN	Region-based Convolutional Neural Networks
VADER	Valence Aware Dictionary for sEntiment Reasoner
TF-IDF	Term Frequency-Inverse Document Frequency
NLP	Natural Language Processing
NLTK	Natural Language Toolkit
DCNN	Deep Convolutional Neural Network
HOG	Histograms of Oriented Gradients
DenseNet	Densely Connected Convolutional Networks
AHRM	Attention-based Heterogeneous Relational Model
GCN	Graph Convolutional Networks
LSTM	Long Short-Term Memory
Word2Vec	Word to Vector
GloVe	Global Vectors for Word Representation
CBOW	Continuous Bag of Words
GRU	Gated Recurrent Unit
VSO	Visual Sentiment Ontology

URL	Uniform Resource Locator
HTML	Hyper-Text Markup Language
SemEval	Semantic Evaluation
Bi-LSTM	bidirectional Long Short-Term Memory
RGB	Red Green Blue (A colouring format)
OCR	Optical Character Recognizer

VI. Conclusion

Finally, this review has reviewed systematically the main themes and results of the literature pertaining to the topic, summarizing in detail the state of affairs in terms of research and practice. It signals major advances that have been achieved in the area while at the same time noting important gaps in the literature that need to be addressed, specifically where empirical data or even mature theoretical models are absent. The integration of current research highlights the multifaceted nature of the topic, demonstrating the interconnection between the different factors that affect outcomes and requiring a multidisciplinary solution. By facilitating collaboration between researchers, practitioners, and policymakers, subsequent studies can draw on the groundwork established in this review to develop more efficient strategies and solutions. As the environment keeps changing, it is crucial that the academic world stays active and responsive to new trends and challenges so that the knowledge acquired is not only theoretical but also applicable. This review is a call to action for continued research work, building on collective knowledge and promoting informed discussion that can spur significant progress in the field.

References

[1]. Sentiment Analysis For Social Media Images By Yilin Wang & Baoxin Li (Nov 2015)
[2]. Reliability And Perceived Value Of Sentiment Analysis For Twitter Data By Jari Jussila, Vilma Vuori, Jussi Okkonen, Nina Helander (June 2017)
[3]. Facial Expression Recognition And Emotion Classification System For Sentiment Analysis By Jayalekshmi J, Tessy Mathew (July 2017)
[4]. Image Sentiment Analysis Using Deep Learning By Namita Mittal, Divya Sharma & Manju Lata Joshi (Dec 2018)
[5]. Twitter Sentiment Analysis Using Natural Language Toolkit And VADER Sentiment By Shihab Elbagir & Jing Yang (Jun 2019)
[6]. Real Time Sentiment Analysis Of Political Twitter Data Using Machine Learning Approach By Joylin Priya Pinto And Vijaya Murari T. (Apr 2019)
[7]. Sentiment Analysis Of Multimodal Twitter Data By Akshi Kumar & Geetanjali Garg (Jan 2019)
[8]. Survey On Visual Sentiment Analysis By Alessandro Ortis, Giovanni Maria Farinella &, Sebastiano Battiato (May 2020)
[9]. Sentiment Analyzer Using Machine Learning By Rakesh Suryawanshi, Akshay Rajput, Parikshit Kokale & Prof. Subodh S. Karve (June 2020)
[10]. Facial Emotion Recognition Using Deep Convolutional Neural Network By Pranav E., Suraj Kamal, Satheesh Chandran C. & Supriya M.H. (Apr 2020)
[11]. Sentribute: Image Sentiment Analysis From A Mid-Level Perspective By Jianbo Yuan, Quanzeng You, Sean Mcdonough, Jiebo Luo (November 2013)
[12]. An Image-Text Consistency Driven Multimodal Sentiment Analysis Approach For Social Media By Ziyuan Zhaoa , Huiying Zhua , Zehao Xuea , Zhao Liua , Jing Tiana , Matthew Chin Heng Chuaa , Maofu Liu (June 2019)
[13]. Sentiment Analysis From Face Expressions Based On Image Processing Using Deep Learning Methods By Aksoy, Selda Güney (May 2022)
[14]. Robust Image Sentiment Analysis Using Progressively Trained And Domain Transferred Deep Networks By Uanzeng You, Jiebo Luo, Hailin Jin, And Jianchao Yang (Septembor 2015)
[15]. Visual Sentiment Analysis By Attending On Local Image Regions By Quanzeng You, Hailin Jin, And Jiebo Luo (February 2017)
[16]. An Overview On Image Sentiment Analysis: Methods, Datasets And Current Challenges By Alessandro Ortisa, Giovanni Maria Farinellab, And Sebastiano Battiatoc (December 2019)
[17]. Cross-Media Learning For Image Sentiment Analysis In The Wild By Lucia Vadicamo, Fabio Carrara, Andrea Cimino, Stefano Cresci, Felice Dell'Orletta, Fabrizio Falchi, And Maurizio Tesconi.(October 2017)
[18]. From Image To Text In Sentiment Analysis Via Regression And Deep Learning By Daniela Onita, Liviu P. Dinu, And Adriana Birlutiu.(September 2019)
[19]. Social Image Sentiment Analysis By Exploiting Multimodal Content And Heterogeneous Relations By Jie Xu, Zhoujun Li, Feiran Huang, Chaozhuo Li, And Philip S. Yu. (September 2020)
[20]. Evaluation Of Deep Learning Techniques In Sentiment Analysis From Twitter Data By Sani Kamış And Dionysis Goularas (August 2019)