Investigating the Impact of Inset Emojis on Images in News Articles

Isaac Muscat

Supervisor: Dr. Dylan Seychell

June 2024

Submitted in partial fulfilment of the requirements for the degree of Bachelor of Science in Information Technology (Honours) (Artificial Intelligence).

L-Università ta' Malta
Faculty of Information & Communication Technology

Abstract

This project employed a mixed method approach to explore the detection of overlaid emojis in images within news articles and evaluate their potential impact on the reader's perceptions. A survey was created to prove whether or not the presence of inset emojis have an influence on human opinion, and once it was proven, a trained object detection model was implemented to be able to detect this form of media bias. Due to a lack of available datasets containing inset images, particularly emojis, a dataset was created on a subset of the COCO dataset as well as the Facebook Emojis Dataset. This dataset was created by placing random emojis from the emoji dataset, performing transformation on them and then placing them onto a background image and saving it as well as its corresponding annotations denoting the emoji locations and type. With the created dataset, two object detection models, YOLOv8 and RetinaNet, were trained to locate emojis within images. The resultant graphs depicted that the inset emoji dataset created had a uniform distribution of emoji types and a normal distribution of sizes, with the centre-point of each emoji being evenly placed around the entire possible area. Furthermore, the Precision-Recall graph displayed how YOLOv8 performed better overall than its RetinaNet counterpart. The survey created consisted of questions asking the respondents about their knowledge on inset emojis in news articles as well as whether or not their feelings change with the presence of inset emoji in an image. The images utilised in the survey were all retrieved from the created dataset. The results found from the survey proved that whilst inset emojis have an influence on human opinion, the change is mainly dependent on the type of emoji used as well as how the emoji correlates to the background image. Moreover, a qualitative interview was conducted with a professional in the media industry. The main topic covered was whether emojis are appropriate to be utilised within news articles, given the serious nature of professional journalism. The results from this research highlighted the influence that inset emojis have on a reader's perception, and so resources including a dataset and object detection models were provided to detect and mitigate this form of media bias.

Acknowledgements

I humbly dedicate this thesis to my family: my late mother, whose strength and resilience inspired me to persevere and pursue my studies despite any hardships I faced; my father, who consistently had my back and supported me throughout the years, even during the toughest of times; and my sister, whose unwavering emotional support has been a constant source of strength for me over the years. I would also like to express my deepest gratitude to Dr. Dylan Seychell for his guidance during the development of this thesis, as well as to the friends I have made throughout this course.

Contents

List of Figures

List of Tables

List of Abbreviations

CNN Convolutional Neural Network.

CSV Comma-Separated Values.

NN Neural Network.

VOC XML Visual Object Classes Extensible Markup Language.

1 Introduction

This project focuses on investigating the impact that inset emojis within news article images have on human opinion, whether their presence alters their perception on the background image or not. A mixed method approach is used where human evaluation is performed to assess whether the hypothesised impact is true, and if proven to be true, the code aspect of the project serves as a form of media bias detection. The code section provides an inset emojis dataset as well as two object detection programs, also called models, trained to locate inset emojis.

Images, which can be described as an exchange of ideas or information through visual representations [1], have become increasingly prevalent in modern society, particularly in media. They serve as a means of communication, capable of evoking emotions in the user, conveying a message to them or even influencing their opinion. Throughout history, images have been heavily utilised to shape cultural narratives or influence human experiences.

Within the landscape of images, there exist two categories: those of which are inset images, and those which are not. An inset image consists of a smaller image being inserted within the border of a larger one. This research focuses on unveiling the influence or change in interpretation created by the addition of an inset image to the original image.

Consider the images below in Figure 1.1, Figure 1.1a depicts a standard landscape, where an individual may evoke some certain emotions or form opinions upon perceiving. However, Figure 1.1b and 1.1c consist of the same base image as the first one combined with some overlaid emoji, in this case one with a crying emoji and the other with a laughing emoji.

(a) Original Image (Non-Inset)

(b) Inset Image with Crying Emoji

(c) Inset Image with Laughing Emoji

Figure 1.1 Base Image vs Inset Images

This dissertation focuses on the analysis of a specific subset of inset images: emojis. Emojis represent a widely recognised form of visual communication in digital contexts. They are graphic symbols which represent facial expressions, abstract concepts, gestures and objects, each of which come with a predefined name [2]. By selecting emojis as the primary focus, this research aims to investigate their unique impact on altering a viewer's perception of the base image it is accompanied with. Emojis serve as an ideal initial test subject for the hypothesis under examination. Their expressive nature allows for nuanced variations in interpretation when inserted onto larger images. Through this investigation, the extent to which the inclusion of emojis influences individuals' perceptions and emotional responses to the overall image is better understood.

1.1 Problem Definition

Media plays a significant role in our lives, and it is crucial for the proper running of a fair democracy. Thanks to media, people from around the world can stay informed on events happening on the other side of the planet instantaneously. Thus, the information that is spread to society should be factual and an accurate retelling of the events that happened. The consequences of misinformation in media could for instance eventually lead to a business' demise, or a political party's following to diminish drastically. Hence, it is crucial for journalists, who are responsible for giving out accurate information about these events, to be as unbiased as possible when disseminating updates on current news to the public. [3, 4]

Studies have been made [5, 6] with respect to identifying the significance of images as a feature in news articles. In one study [5], Hu et al. researched specific features within news articles that can influence the user's engagement positively. The results, denoted by linear regression graphs, proved that images have a major impact on the user's behaviours and interest when reading articles. Similarly, in another study [6], Sargent investigated the influence of images on a news article's reading time. The paper concluded that images on the front page of a news story, or nowadays, a thumbnail, are more likely to appeal to readers.

However, the relationship between journalism and images in news articles could lead to a potential problem: the risk of journalists utilising them incorrectly for misinformation purposes. Whilst images certainly enhance the user's engagement time with a news article, they can also be used to distort facts. Therefore, it is the journalist's responsibility to accurately retell events, rather than prioritise sensationalism to fit a narrative. Otherwise, such misinformation could perpetuate biases and misdirect the public. [7]

To address this problem, this research investigates the use of inset images, specifically emojis, within news articles, and how journalists may utilise them to influence the interpretation of the original base image. Additionally, this necessitates exploring practices for the mitigation of such risks associated with the misuse of visuals for sensationalism or misinformation purposes. This research seeks to contribute to the ongoing discourse on media ethics and journalistic integrity, advocating for a more transparent approach to visual narration in journalism.

1.2 Motivation

In recent years, media bias analysis has undergone major changes mainly due to the sudden increase in information dissemination. In response to this sudden surge, some researchers [8–10] have started to integrate automated methods to detect types of bias in media. Whilst most of these tools are focused on the textual component of news articles, some noteworthy projects [3, 11] have been made to address bias in images as well.

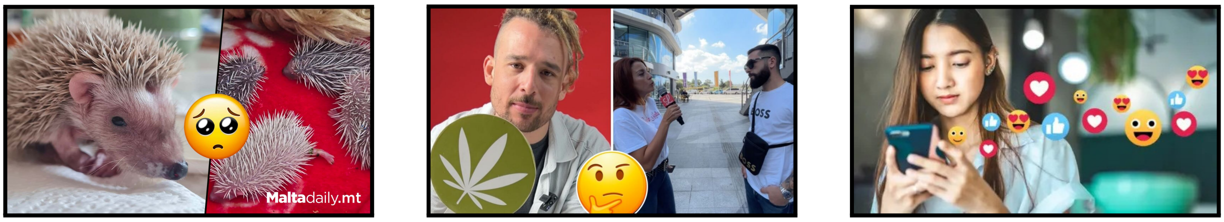
Despite these developments, challenges still persist in image bias detection, particularly in cases where inset images are present. These challenges originate from the unique nature of images, which are not as easily analysed for bias as textual content. Hence, this type of automated bias detection must cater for the intricacies found in image-based media. Moreover, this research has opened up possibilities for interdisciplinary collaboration between researchers of multiple fields. The fields of Artificial Intelligence, Media Studies and Communication Research all foster a holistic understanding of the presence of bias in information dissemination.

Whilst efforts have been made in automating bias detection in image-related media, the complexities introduced by such inset images, particularly emojis, underscore the need for further research in the field. As technology continues to advance, a collaborative effort remains crucial to refine automated bias detection methods and ensure a comprehensive analysis of such bias in diverse media formats.

1.3 Problem Challenges

The decomposition of inset emojis within news articles necessitates the application of multiple computer vision techniques. However, the execution of these techniques is contingent upon the availability of a suitable dataset. While numerous datasets exist for normal images or emojis, the absence of dedicated datasets for inset images, especially those with emojis overlaid on other images, poses a significant challenge. This gap in available datasets impedes the development of models capable of detecting inset images and predicting the impact they may have on readers' opinions.

Despite the increasing prevalence of inset emojis in news articles, as can be seen in Figure 1.2, whose images are taken from the popular local Maltese news websites Maltadaily, LovinMalta as well as in international news articles with an example image taken from BBC's website, this area remains largely unexplored. In response to this, our study is motivated by a dual objective: the creation of a dataset to facilitate advanced object detection techniques and human evaluation to research the effects inset images may have on human opinion.



(a) Maltadaily Inset Emoji Example 1 [12]

(b) LovinMalta Inset Emoji Example [13]

(c) BBC Inset Emoji Example [14]

Figure 1.2 Local and International Newsrooms Article Thumbnails using Inset Emoji Examples

1.4 Aims and Objectives

The aim of this project is to develop a model capable of detecting emojis, whilst also exploring the predicted impact inset emojis have on human opinion. This will require an object detection model capable of detecting emojis inset onto a larger background image, whilst also performing human evaluation regarding how individual's opinions on an image may differ based on the presence of inset emojis and their characteristics. For such a task to be achieved, the main objectives are to:

- **Objective 1 (O1)** - Create a dataset to train models on, consisting of random images retrieved from an open-source dataset and emojis, placed onto the larger image to create inset images.
- **Objective 2 (O2)** - Train an object detection model to predict the locations of the inset emojis within the whole image, as well as its respective emoji type.
- **Objective 3 (O3)** - Perform Human Evaluation on the dataset, composing of quantitative and qualitative analysis, analysing how an individual's opinion changes depending on the presence of an emoji in an image. The images to be used for this section are retrieved from the dataset created.

Overall, this project comprises two separate components: the code-related aspect of the project, consisting of the dataset of inset images and the object detection model as well as the Human Evaluation part, which will study the effects of inset emojis on human image interpretation. If the hypothesis that inset emojis have an impact on reader opinion is proven to be true by the human evaluation results, the object detection model can be used to detect inset emojis in media and assess the potential influence they may have on the audience.

1.5 Proposed Solution

This study proposes a new dataset that consists of emojis of different sizes and rotations inset randomly onto a set of images to provide the research community with a useful resource. Apart from a dataset, this project also trains object detection models to test the performance of two of the most highly utilised object detection models available on the dataset and evaluate their performance. Apart from this, human evaluation is conducted to conclude whether or not the inclusion of an inset emoji has an influence on the reader's perception on the image, and if it does, how so. This section of the research will consist of a survey being handed out to multiple people as well as an interview with a professional in the media industry to gather insights into the utilisation of inset emojis.

1.6 Organisation of Document

Chapter 2 provides background information on media bias, datasets, machine learning methods, and the literature review. Chapter 3 details specification and design, including an overall explanation of how the code section of the project was developed. Chapter 4 covers implementation and design reasoning, explaining certain choices made by using block diagrams and other types of figures. Chapter 5 presents project evaluation and the results concluded from it. Finally, Chapter 6 offers a brief project overview, conclusive remarks made as well as ideas for future work.

2 Background

This chapter delves into the research area of inset emojis by examining relevant literature, offering insights relevant to the project's development. The background section explores media bias, datasets, and machine learning principles, while the literature review covers research in image classification, object detection, and similar research within the emoji field. Finally, some studies related to the effects of emojis on human emotion are also discussed.

2.1 Media Bias

Media bias is the systemic prejudice in the dissemination of information by media outlets, influencing narratives across various subjects such as politics, economics, and culture. This bias significantly shapes the public's perceptions and opinions, becoming a crucial aspect to explore.

According to a study made by Puglisi et al. [15], there exist two types of media bias: distortion and filtering. Distortion involves the presentation of information in news reports while omitting relevant facts or failing to cover the full story. This can lead to a skewed understanding of events by the audience. For instance, within the realm of images, consider a news story about a political rally where supporters of Candidate A outnumber supporters of Candidate B. However, a photograph accompanying the story may be cropped in such a way that it only shows the smaller crowd of Candidate B's supporters, giving the false impression that their turnout was comparable to or even greater than that of Candidate A. This selective use of images can distort the perception of the event and influence viewers' opinions in favor of Candidate B, despite the actual turnout favoring Candidate A. As can be seen in Figure 2.1, an example image taken from research by Fanfani et al. [16] regarding the automated detection of cropping in images depicts an illustration of ambiguity resulting from image cropping. In this image, an Iraqi soldier is surrendering to the U.S. army. The difference in interpretation can be noted when cropped as can be seen in the left part of the image and the right when compared to the original image in the center.

Figure 2.1 Media Image Distortion - Left: Gun pointed at an Iraqi soldier. Right: US soldier providing a drink to an Iraqi soldier. (Source: [16])

The second type of media bias, filtering, occurs when news reports summarise events and by doing so decide which facts to include and which to not. This process allows a news outlet to frame the sequence of events in a way to fit its ideologies or agenda, leading to a biased portrayal of the actual truth. For example, consider a protest rally where the majority of the protesters are participating in a peaceful demonstration. However, a few other protesters partake in a more violent protest. A media outlet, depending on their stance, may opt to only mention the violent protesters, giving the impression that the protest as a whole was chaotic. On the other hand, an opposing outlet may only focus on the peaceful demonstrators, ignoring the violent protesters and presenting the entire event as a peaceful one. [17]

In the realm of news articles, media bias extends to the visual domain. Media bias in images involves the deliberate selection and presentation of images to potentially distort the true facts intended by the article. This can arise from choices related to the image selection, the angle it was captured from, or any editing, such as cropping. Portraying a country's cleanliness with a front image of a littered dirty street when compared to a newly renovated building can drastically alter the overall perception of the city depicted by the article.

For example, consider Figure 2.2 below, both images were retrieved from the Times of Malta website and depict an area in the capital city of Malta, Valletta. The image on the left displays a dirty set of steps within Valletta whereas the image on the right shows a cleaner side to the city. These images were specifically chosen to better fit the narrative that the article is covering, hence possibly influencing the opinions on the city's cleanliness through visuals.

Dirty Street in Valletta [18]

Tritoni Fountain in Valletta [19]

Figure 2.2 Image Selection Comparison

Inset images, on the other hand, compound the potential bias by combining two separate images. This technique involves the strategic use of smaller images within a larger context, often placed alongside the main image to emphasise specific aspects of a story. These insets can subtly sway the public's interpretation of the text they are reading. For instance, placing an angry emoji next to an image of a parliament building could influence readers' perceptions of the overall news story, even if the text itself does not necessarily reflect the emotion evoked by the inset image.

2.2 Datasets Used

For model training purposes, this project required the use of two datasets, one consisting of general images with no particular theme, to keep model generalisability as high as possible, and the other consisting of a set of emojis to be placed onto the general images. The two datasets used were the COCO dataset accompanied by a subset of the Facebook Emojis dataset.

2.2.1 COCO Dataset

The COCO dataset is a widely used dataset in the field of computer vision and object detection. Created by Microsoft Research, COCO is designed to cater for the limitations of dataset that came before it by providing a diverse range of images. COCO's main advantage resides in its diverse environments containing multiple different types of objects. Over the years, COCO has become a key resource for Computer Vision related projects, used in classification, object detection and even segmentation. [20]

2.2.2 Facebook Emojis Dataset

Facebook emojis act as a powerful communication tool with regards to expressing emotion and reactions. Offering a vast and diverse selection, these emojis allow users to convey their sentiment in relation to some message or post online through a small image, which can depict different emotions, such as a smiling face, an angry face or even a thumbs up to show agreement. For this project, a subset of the available dataset of Facebook emojis was selected, these being Angry, Care, Haha, Like, Love, Sad and Wow as can be seen in Figure 2.3 below. The main reason why this subset was selected was due to literature about Facebook Emojis mainly focusing on the react emojis on posts, these being the seven emojis selected for the dataset. This approach allowed for an accurate evaluation of the impact of emojis on user perception and engagement, shedding light on their role in shaping online discourse and social interactions. [21]

Figure 2.3 Facebook Emojis Used retrieved from IconScout [21]

2.3 Machine Learning Methods

2.3.1 Neural Networks

Neural Networks (NNs), part of a subset of Artificial Intelligence known as Machine Learning, consist of artificial neurons utilised to process input data through weighted transformations across multiple layers, adjusting weights to minimise the output error throughout. Traditional NNs have waned in use for high-accuracy tasks, giving way to Deep Neural Networks (DNNs) which are beneficial for representing complex data such as images and audio [22, 23]. DNNs have multiple subtypes built for their own unique purposes, these include Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory Networks (LSTMs), and Generative Adversarial Networks (GANs). CNNs notably excel in image processing by preserving spatial relationships through convolutional layers, enabling effective feature extraction for image classification and object detection tasks, including the ability to recognise patterns regardless of their location, also known as translation invariance [24].

2.3.2 Training a Network

Training a Convolutional Neural Network involves multiple steps for a model to learn patterns and features within some inputted data, as outlined by a paper by Liu et al.

[25]. The first step is to prepare the training set, where images are preprocessed to be standardised in the appropriate formats and their pixel values are normalised. Following this, the neural network's architecture is defined, comprising multiple different layers. The next step involves creating a loss function to measure the difference between the predicted result and the ground truth. The model then adjusts its weights during each training step to minimise the loss. In this paper, models were used with pre-trained weights, however these models were then trained on the datasets used.

2.3.3 Object Detection

Object Detection is a type of algorithm in the field of computer vision that involves identifying and locating an object or multiple objects within an image. Its predecessor, classification, is responsible for assigning a label to the entire image, however object detection acts as an advancement to this by locating the object to be labeled. It places a bounding box around the predicted position of the object and assigns it a specific class label.

To train an object detection model, the training dataset must first be created. In this case, the dataset consisted of the images containing the objects to be detected, a file containing the predefined list of labels to be detected, and another file which contains a list of the images with the respective coordinates and labels of the objects to be detected from them. The training set's format is dependent on the type of model being used, some formats include but are not limited to: COCO, Pascal VOC and YOLO. Once the model was trained and tested, it could then be used for detecting objects in unseen examples. [26, 27]

2.4 Metrics

The following equations are a list of metrics that were utilised during the evaluation process.

Precision is the metric that defines how many items retrieved by the model were relevant.

$$\textbf{Precision} = \frac{\text{True Positive (TP)}}{\text{True Positive (TP) + False Positive (FP)}} \tag{2.1}$$

Recall is the metric that describes how many of the relevant items were successfully retrieved.

$$\textbf{Recall} = \frac{\text{True Positive (TP)}}{\text{True Positive (TP) + False Negative (FN)}} \tag{2.2}$$

The F1-Score is the metric that serves as a harmonic mean between precision and recall.

$$\textbf{F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision + Recall}} \tag{2.3}$$

A confusion matrix is a tool used to assess the quality of predictions against an expected ground truth.

$$\textbf{Confusion Matrix} = \begin{array}{r|cc} & \textbf{Predicted Positive} & \textbf{Predicted Negative} \\ \hline \textbf{Actual Positive} & \text{True Positive} & \text{False Negative} \\ \textbf{Actual Negative} & \text{False Positive} & \text{True Negative} \end{array} \tag{2.4}$$

Mean Average Precision (mAP) is a metric that considers both the precision and recall values across multiple thresholds.

$$\textbf{mAP} = \frac{1}{N} \sum_{i=1}^{N} \text{AP}_i \tag{2.5}$$

N is the number of classes, and AP_i is the Average Precision (AP) for class i. Average Precision (AP) for a single class is the area under its corresponding precision-recall curve.

2.5 Literature Review

In the field of computer vision, classification stands as one of the most fundamental cornerstones, pivotal in enabling models to assign a label or labels to a given image by extracting its features and comparing them to a predefined set of labels. Similar to the training of any neural network, the training process in Classification involves the model learning the given labeled data, so that the model can then accurately label unseen data by recognising patterns within it that are indicative of specific classes. Classification also serves as a basis for more complex tasks in computer vision, such as object detection and image segmentation. [28]

Object detection, another computer vision task which builds upon the advancements of classification is tasked with locating objects within a given image or video. However, unlike image classification, object detection is capable of identifying the location of an object within the image via a bounding box, which is then assigned a label. This functionality is crucial for a wide range of applications, including surveillance [29], autonomous navigation [30], and augmented reality [31]. Over the past few decades, object detection has had many advancements for it to be able to produce the results possible today. The classical approaches, such as Haar Cascades

[32] and Histogram of Oriented Gradients [33], laid the foundations for object features however fell short with the problems of scale variations and occlusions. To solve this, the shift towards Convolutional Neural Networks (CNNs) greatly increased the overall precision and accuracy possible. CNNs are adept at learning representations in visual features, enabling them to effectively handle scale variations, occlusions, and complex object arrangements. Thanks to these advancements, object detection's performance has seen significant improvements across various domains, making it one of the most crucial components of modern computer vision systems.

In addition to the technical aspect, understanding the influence of emojis on human perception and communication is also relevant to this project. Emojis, widely-utilised in modern digital conversations, are a useful tool when it comes to clarifying the mood and tone of a message. In marketing for example, emojis are leveraged to reinforce brand personality and engage audiences [34, 35]. Research in this area provides insights into the intersection of technology, communication, and human psychology, informing more effective strategies for online interaction [36–38].

2.5.1 Image Classification

The landscape of image classification shifted completely and gained significant traction with the development of CNNs, pioneered by LeCun et al. in 1998 [39]. Using CNNs, image classification was capable of learning features from raw pixels data, enabling more accurate predictions than was previously possible.

Classification for Emojis

Over the past few years, with the major advancements in image classification, multiple projects related to emoji classification have been made [40–42]. Hossain et al. [40] researched the classification of hand-drawn emojis and how they can be classified into categories of emojis. The researchers opted to use a CNN-based approach, training a model architecture consisting of four Convolutional layers each followed by a Max Pooling layer. A custom dataset was used to train the classification model, consisting of 4000 images, 500 images per label.

In another study, Bala et al. [41] developed a project which focuses on classifying emojis using a Random Forest Classifier (RFC). The emojis, which were retrieved from multiple sources including Apple and Facebook emojis, were classified into six labels, these being: ”Positive”, ”Negative”, ”Happy”, ”Angry”, ”Sad” and ”Love”. The classification process involved training the RFC using the dataset organised into two separate scenarios. In the first scenario, each emoji instance was processed based on the aforementioned six classes, providing individual assessments for each emotion conveyed by the emojis. In contrast, the second scenario simplified this task by limiting the categories to two classes, one representing positive sentiments and the other representing negative sentiments, while also considering both the emoji and the accompanying text in each assessment.

Bharati et al. [42] researched the classification of emojis using CNNs, presenting an innovative approach to automatically assign a facial expression to a particular emoji. Two machine learning models were created, one for face recognition, which is based on the Fisher-face algorithm, whilst the other utilised the VGGNet Network. The models were trained on Tensorflow using a dataset retrieved from Kaggle, comprising over 12,000 emojis categorised into twenty classes.

2.5.2 Object Detection

Object detection has evolved drastically over the past decade. Traditional approaches in object detection relied on sliding window techniques and handcrafted features whereas this all changed with the advancement of CNNs.

Girshick et al. [43] provided further advancements to this field with the release of Region-Based CNN which improved the accuracy of object detection by proposing regions of interest, however still suffering from computational inefficiencies. In the following years, many other advancements were made. Faster R-CNN by Ren et al. [44] introduced Region Proposal Networks, used to streamline region proposal generation. Liu et al. introduced Single Shot Multibox Detector [45] shortly followed by Redmon et al's You Only Look Once model [46]. These new models marked a shift towards real-time object detection by predicting bounding boxes alongside their corresponding labels with a confidence percentage simultaneously.

Faster R-CNN

This object detection model, released in 2015, utilises Region Proposal Networks. The purpose of the Region Proposal Network is to efficiently propose candidate bounding boxes in an image where objects might be located. Anchor boxes, which are pre-defined bounding boxes, are used, from which the model predicts whether an anchor box fits the label or not. Faster R-CNN achieves a balance between accuracy and efficiency by using its two-stage architecture, setting the stage for subsequent advancements [44].

Mask R-CNN

Mask R-CNN is an extension to Faster R-CNN. Released in 2017, it introduced instance segmentation capabilities. This means that whilst still predicting bounding boxes, Mask

R-CNN is also capable of providing pixel-wise segmentation of objects within images [47].

RetinaNet

RetinaNet, released in 2017, addressed a common problem faced by many object detectors, this being class imbalance. This model introduced focal loss, a function that prioritises examples which are hard to classify by handling the imbalance between foreground and background classes. By mitigating the background class, RetinaNet managed to achieve remarkable results in object detection tasks [48].

EfficientDet

Introduced in 2019, EfficientDet optimises both accuracy and model size. Its model utilises a scaling method to ensure its deployability on devices with low computational power. By addressing this trade off, EfficientDet has become a crucial model in the landscape of object detection models [49].

YOLOv8

YOLOv8, the second latest iteration in the YOLO series, is one of the most significant advancements in object detection. Released in 2021, YOLOv8 builds upon its predecessors, incorporating improved architecture and training strategies. This model focuses on maintaining real-time functionality whilst enhancing accuracy [50].

2.5.3 Logo Detection

A subsection of object detection similar to inset image detection is logo detection. Researchers in the field have employed multiple methodologies and written about the challenges faced in such a task.

Research by Dey et al. [51] focused on the detection of stamps and logos within document images using an outlier detection approach. Essentially, the method utilised scalar colour feature extraction, followed by segmentation. Hence, the foreground pixels were isolated using a process called Principal Components Analysis to distinguish the main colour. The connected parts of the images were then grouped together into processing units based on a predefined height and width criteria. Feature extraction was performed on these units, by considering their line thickness, height, position and pixel density. Outlier techniques were used too, by first selecting all units based on particular features and then testing them on a list of criteria. The method was tested on a dataset and metrics such as precision and recall were used for evaluation.

The study provided insights into the complexities of logo detection, emphasising the need for development in handling overlapping elements in document images.

A similar research done by Lu et al. [52] explored logo detection, however this time utilising Convolutional Neural Networks. The authors incorporated a Global Pyramid Text Attention module and a Pyramid Text Attention module to enhance both logo and text detection within the given document images. The Global module suppressed background interference, generating saliency maps, which are maps that highlight the most important regions within an image, whilst the Pyramid module addressed logo detection by aggregating features for local detail and semantic information. The study emphasised the crucial post-processing step of eliminating any redundant bounding boxes by using Non-Maximum Suppression. This was done by ensuring that the detected object is associated with the highest-scoring bounding box. The model was tested on datasets including FlickrLogos-32 and TopLogos-10 and demonstrated a superior performance to other models that came before it. This paper showcased the significance of advanced CNN architectures and their post-processing techniques in achieving accurate and reliable logo detection.

Rezkiani et al. [53] researched Logo Detection but with YOLO as an architecture. They presented an approach utilising the YOLOv4 algorithm for accurate logo detection in diploma images. The images were gathered from multiple universities and were converted into the same format to be able to perform data augmentation on. After training, the model was then capable of dividing a given diploma image into a grid and accurately predicting the location of the logo. The study evaluated the model's performance by using a confusion matrix and metrics such as precision, recall, F1-Score as well as the Mean Average Precision values. The result demonstrated YOLO's effectiveness in logo detection and classification, achieving very high accuracy results, showcasing its practical applicability in object detection scenarios.

2.5.4 Emoji Detection

The realm of emoji detection, particularly those embedded onto a larger image remains to be explored profoundly. Singh et al. conducted a project related to emoji detection [54], where they developed a dataset consisting of hand-drawn emojis and trained an object detection model on them. The architecture for the model consisted of three Convolutional layers, each followed by a pooling layer respectively. The dataset was split into eight categories, with the total number of images amounting to 2000 images. Following hyper-parameter optimisation and testing, the model achieved an accuracy of 63.56%.

2.5.5 Emojis Influence on Human Emotion

Apart from analysing the advancements of technology with regards to object detection models, this study also pertains to uncovering the effects inset emojis may have on the overall interpretation of an image. Some noteworthy studies have been made [55–59] to study the effects of these emojis, and how they are used to convey or alter the interpretation of the accompanying text.

In a study by Tian et al. [55], the dynamic of Facebook reactions and emoji usage were inspected across diverse cultural and linguistic contexts. The dataset utilised consisted of 21,000 posts retrieved from public media pages across four countries, which were used to scrutinise the distribution of Facebook reactions and the prevalent usage of emojis in user comments. One notable difference found was the variations in emoji reaction usage among nations, highlighting cultural influences. Moreover, this study also investigated the nuanced relationships between reactions and emoji sentiment, giving light to the importance of considering the accompanying context when interpreting them online.

Additionally, a study by Das et al. [56] researched the influence of emojis, specifically positive ones, in marketing. An example of a positive emoji for instance would be a smiley face, as it conveys a positive emotion. The study concluded that the utilisation of emojis in marketing, particularly positive ones, are likely to enhance the positive effect on customers to entice them to purchase some product. However, the effectiveness is also dependent on other characteristics, mainly the type of product being advertised. The research found that emojis have a greater influence on the customer when paired with entertainment related products, whereas a significant reduction in engagement was noted when paired with utilitarian focused products such as furniture or tools. Thus, the influence of emojis on the users opinion is not only dependent on the emoji itself, but also what it is being coupled with.

In a similar study by Ko et al. [57], research was done on the influence emojis may have on user engagement in brand-related user generated content. The results of the research suggested that the incorporation of emojis within User Generated Content (UGC) could boost the overall Consumer Engagement (CE). However, they also noted that an excessive amount of emoji usage may not necessarily increase this, rather it may sometimes cause a negative effect. Specifically, emotional emojis tended to enhance CE when paired with positively skewed text. Moreover, positive emotional emojis demonstrated more impact as opposed to general ones in terms of CE. Hence, this research provided insights with regards to the possible effects emojis may have in marketing and branding, and how the emoji type as well as the text it is paired with all have an effect on the overall CE.

Research by Wibowo et al. [58] explored the effect emojis have when present in text messages, and how they affect the reader's perception on the message sender. The study involved 48 undergraduate students split into three groups, one which were given a smiling emoji, one with an unamused emoji and the other with no emojis at all. The results indicated that the presence of a smiling emoji influenced the reader's perceptions of the sender to be more sincere and friendly, similar to the ones without emojis, indicating the importance of the verbal content itself. However, the presence of a frowning emoji performed the opposite of this, influencing the reader's perceptions of the sender to have more of an unfriendly attitude. The results and conclusion of this study suggested that while the presence and use of emojis have an influence on the perceptions of the user, the content used and the context are also very important. Furthermore, the interpretation of emojis could be different dependent on the demographics of the user, be it age group or even gender.

Finally, in a recent study by Maiberger et al. [59], the impact of facial emojis in electronic word of mouth (eWOM) was researched based on the Emotions as Social Information (EASI) theory. This model suggests that the way we express emotions can influence others by giving them insights about the situation. The study found that facial emojis influence persuasion mainly through emotional arousal and perceived ambiguity. While emotional arousal typically increases persuasion, the effect of ambiguity depends on the functionality of emojis in electronic word-of-mouth (eWOM): when emojis substitute verbal expressions, ambiguity rises, diminishing persuasion; conversely, when emojis reinforce verbal expressions, ambiguity decreases, increasing persuasion. These two influences collectively determine the overall effect emojis have on persuasion in eWOM.

3 Specification and Design

This section provides an overall explanation of the design of the project made, explaining each step performed through the help of images and diagrams. The chapter consists of two subsections: the dataset creation part which covers the dataset's architecture and how it was created and the object detection models part, which explains the choice of models as well as the training model process. Each subsection will explain the process on how such tasks were carried out as well as the reasoning behind the design choices made.

3.1 Dataset Creation

The dataset comprises two sections: the classification section and the object detection section. These two subsections were implemented to provide users exploring the field of inset emojis with as many resources as possible to further expand or develop classification and object detection models.

3.1.1 Organisation of Dataset

Classification

The classification subsection is composed of three subfolders, a Training, Validation and Testing subfolder, each of which contains a folder containing normal images and the other containing inset emoji images. It is important to note that the classification subfolder consists of 1500 non-inset images and 4500 inset images. For training a classification model, it is recommended to split the inset and non-inset training sets equally to avoid bias, for instance by taking 1500 images from the non-inset and inset images sets each.

Object Detection

The object detection subsection is composed of three subfolders: YOLO, VOC XML and CSV. Each subfolder is split into a Training, Validation and Testing set, which contain a folder of images and their respective annotations. In the YOLO subfolder the annotations consist of .txt files, the VOC XML subfolder contains .xml files, one file for each image. The CSV subfolder contains one .csv file per subset of images, for instance one file for the training set.

3.1.2 Creation Process

Requirements

The creation of the dataset required two datasets which were previously mentioned. The COCO dataset was used to serve as the background images, and a subset of the Facebook Emoji images was used to serve as the inset images. Due to the size of the COCO dataset, only a subset of these images were chosen to keep the dataset's size reasonable while still being feasible for training models. Thus, 1500 images were selected, with which a total dataset of 6000 images could be created.

Creating an Inset Image

When creating an inset image, three subsets of images were made: those with one, two and three inset emojis. The emoji to be placed was randomly selected from the subset, to which four transformations were performed:

- Rotation - The emoji was transformed on a singular central point by any angle on the z-axis.
- Stretching - The emoji was transformed on either the x-axis or the y-axis by expanding or compressing the dimensions of the image.
- Scaling - Similar to stretching, however the scaling factor is equal for both dimensions of the image. A shrinking or enlargement effect was seen, where the original proportions of the image were retained.
- Translation - The emoji was moved from one pixel position to another without changing its size or rotation.

Furthermore, checks were made to ensure that all emojis are mostly visible and not completely cut off or covered by other emojis.

Saving an Inset Image

After placing the emoji on the image, its characteristics were saved to be used in the subsequent section which involves annotation conversion and saving. The characteristics saved are:

- Coordinates - The x and y coordinates were saved, with x1 and y1 serving as the starting point of the image, and x2 and y2 marking the ending coordinate.
- Emoji Type - The emoji type was saved to separate the different types of emojis used in the image, enabling easy categorisation and detection.

- Image Name - The image name was stored to provide a unique identifier for the image.

For the image name, the format of **Emojis_*{numOfEmojis}*_*{image}*.jpeg** was utilised. **numOfEmojis** represents the number of emojis inset into the image. For instance, an image with three inset emojis would have 3 as **numOfEmojis**. Furthermore, **image** is the name of the original image retrieved from the COCO dataset, such as 'Image1000'. Consider the image in Figure 3.1, an example name for the image would be 'Emojis_3_Image10.jpeg'.

Figure 3.1 Example Image Generated

After the image's characteristics were retrieved, there were three formats in which the image annotation information was saved:

- **YOLO**
- **VOC XML**
- **CSV**

Each of these annotation types required the same information, however in different orders or formats. With respect to YOLO, the x2 and y2 coordinates are replaced by the width and height of the emoji, simply calculated by subtracting the original x2 and y2 by x1 and y1 as can be seen in (3.1) and (3.2). Also, the x1 and x2 values are replaced by the center point of the image, so for instance x1 is replaced by x1 added with half of the width of the emoji. The four values are then normalised by dividing them by the background image's respective width or height, depending on whether the coordinate is of the x-axis or the y-axis. The four coordinates can be better visualised in the diagram in Figure 3.2 below.

$$\text{width} = |x_2 - x_1| \tag{3.1}$$

$$\text{height} = |y_2 - y_1| \tag{3.2}$$

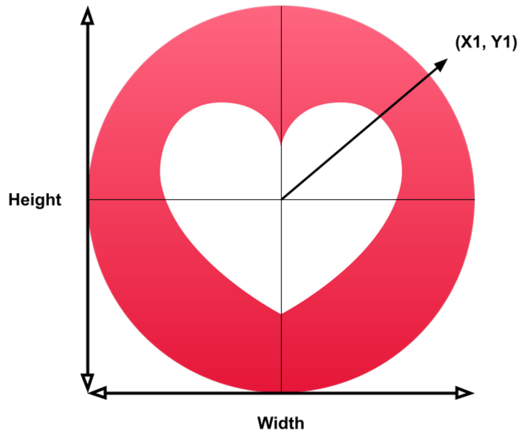


Figure 3.2 YOLO Annotation Coordinate Positions

The YOLO information was then saved into a .txt file, one for each image saved. Also, for cases in which multiple emojis were present in the image, a new record of information was simply added to the pre-existing .txt file. Both the format used and an example file based on the exemplary image provided in Figure 3.1 can be seen in Table 3.1 below. Note that in the .txt file itself, the values were simply separated by a comma, the table was only used for presentation purposes.

Table 3.1 YOLO Annotation Example

Emoji Type	X1	Y1	Width	Height
1	0.682	0.536	0.186	0.248
4	0.716	0.709	0.192	0.256
5	0.108	0.742	0.216	0.329

The VOC XML information consists of the original x1, y1, x2 and y2 coordinates unlike the YOLO format. This format contains information about the image as a whole such as the image name, width and height, followed by the information about each annotation. For each image, an XML file was created, thus if an image consisted of multiple emojis, information was simply appended to the file. Moreover, both YOLO and VOC XML annotation files retained the same naming convention as that of the images. The format used may be better visualised using the example depicted in Figure 3.3 below. An example of the XML file generated can be seen in Listing A.1 in the Appendix.

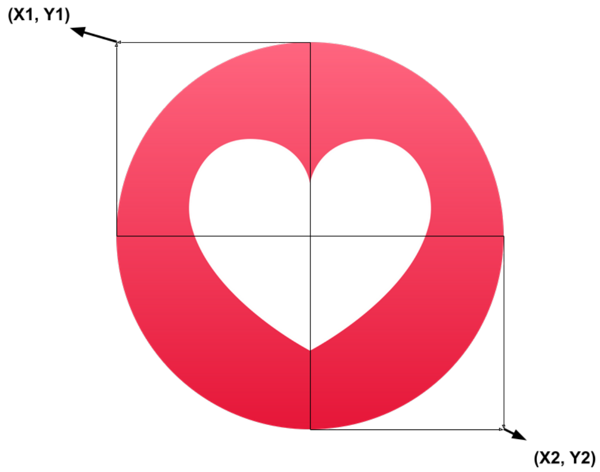


Figure 3.3 VOC XML Annotation Coordinate Positions

The CSV information consists of the same information as that of the VOC XML files. However, rather than each image having a corresponding file, one CSV file was created to store all of the information of every image created within the respective set. An example of the format can be better visualised in Table 3.2 below.

Table 3.2 CSV Annotation Example

Image	X1	Y1	X2	Y2	Emoji Type
Emojis_3_Image311.jpeg	377	198	496	317	care
Emojis_3_Image311.jpeg	397	279	520	402	love
Emojis_3_Image311.jpeg	0	277	138	435	sad

The images as well as their corresponding files were saved into one of three respective subfolders, these being the Training, Validation and Testing folders. These three subfolders can be found in the Object_Detection folder. This split helped ease the training and testing of object detection models in the subsequent section of the project. The split ratio chosen was 60-20-20, based on another computer vision related task which used the same split ratio [60].

3.2 Object Detection Models

Two models were implemented to detect emojis using the dataset created for training. The models selected were YOLOv8 and RetinaNet, due to them being two of the best models currently available for object detection and working well with objects of different scales and sizes [61, 62].

3.2.1 Training Models

Each model required setting up and installing the required packages to function properly. After setting up each model, the models were trained on the training set, consisting of 3000 images. The training for both models was set to ten epochs, to avoid underfitting or overfitting the model on the training set.

3.2.2 Testing Models

Following training, the models were then validated by using the Validation set, and tested using the Testing set. The purpose of this subsection was to evaluate the performance of the trained models on unseen examples.

3.2.3 Evaluation

After testing, graphs were displayed and saved to be able to visualise the performance of the model generated. For YOLOv8's model, by performing the test function multiple graphs were saved, these being:

- Confusion Matrix - A matrix depicting the number of True Positives, True Negatives, False Positives and False Negatives.
- F1-Curve - A graph depicting the harmonic mean of precision and recall changes across different classification thresholds.
- PR-Curve - A graph representing the trade-off between precision and recall, with precision on the y-axis and recall on the x-axis.

For RetinaNet's model, a PR-Curve was generated, as well as a Mean Average Precision graph, which is a bar graph that depicts the performance of the object detection model across the different labels.

4 Implementation

This section delves into greater detail with respect to both the dataset creation as well as the object detection model training, whilst also explaining other files provided within this project. A link to the code repository can be found in Appendix C.

4.1 Dataset Creation

4.1.1 Packages Required

In the dataset creation notebook, multiple packages were required to be installed and imported to be able to carry out the process.

Image Loading and Handling

To be able to load images and perform transformations on them the following packages were required.

- **cv2** - OpenCV is a library of functions mainly related to computer vision. In this project, this library was used for loading and saving images.
- **tensorflow** - TensorFlow is an open-source machine learning framework. This package was utilised for transformations on images such as resizing the image.
- **tensorflow_addons** - TensorFlow Addons is a repository of functions not available in core Tensorflow. This package provided some additional functionalities such as the rotation function which was used for data augmentation purposes when creating an inset emoji.

File Structure and Saving

The following packages were used for saving and managing information in specified formats:

- **csv** - This module contains functions to read and write tabular data in CSV (Comma Separated Values) format. This was used when generating the CSV annotations for the inset images.
- **xml.dom.minidom** - This module provides a minimal implementation of the Document Object Model (DOM) interface, primarily for XML processing. The parseString() function of this module was utilised for presentation purposes when saving the VOC XML annotations.

- **xml.etree.ElementTree** - This module provides a simple and efficient way to parse and manipulate XML data. XML code was generated using this module to be able to save the annotations in VOC XML format.

4.1.2 Dataset Creation Walkthrough

After the necessary libraries and packages were installed and imported for the notebook, the required datasets were appropriately set up. The user is given the option to select which dataset to use for background images, whether it is the COCO dataset or some alternative dataset. In the case of the COCO dataset, due to its size, 1500 images were retrieved from the dataset and saved temporarily in a dictionary. After creating the dictionary of background images, a similar process was performed for creating a dictionary of emoji images.

The next step involved defining functions to create and save the dataset created appropriately alongside each image's corresponding annotations into specific formats. These functions are responsible for opening or creating a file to save the annotation data into. If the file already exists, information is simply appended to it. The algorithm used to generate inset images can be seen in Algorithm 1.

Algorithm 1 Inset Emoji Generator Pseudocode

```
procedure EmojiInsetFunc(currentImages, emojiImages, numOfEmojis,
        directory, classificationDirectory)
    for image in currentImages do
        if numOfEmojis ≠ 0 then
            for emojiCounter in range(numOfEmojis) do
                Select a random emoji from emojiImages
                Rotate and Stretch the emoji
                while image position is not valid do
                    Resize the emoji
                    Calculate the area taken up by the emoji on the image
                    if emoji overlap size does not exceed limit then
                        Image position is valid
                    end if
                end while
                Generate annotation data for the emoji
                Save the annotation data
            end for
            Save the inset image
        else
            Save the non-inset image
        end if
    end for
end procedure
```

The main emoji inset image creation function required five parameters: the dictionary of background images, the dictionary of emoji images, the number of emojis to inset into one image, the directory to save the object detection annotations and images and the directory to save the classification images. The function first loops through all of the background images in the dictionary. In each iteration of the loop, another loop was performed for the number of the emojis to inset onto the background image.

For each iteration of the inner loop, a random emoji was selected. The emoji was transformed randomly in terms of rotation, scaling, stretching and translation. With respect to rotation, the angle at which it could be rotated was made to be between 0 and 360. For stretching, the emoji was allowed to be stretched by a maximum of 1.24 times its original size, the reasoning behind this was that in a similar paper by Niu et al. [63], where 1.24 was the maximum stretching factor utilised. With respect to scaling, a paper by Wang et al. [64] experimented with how recognisable a background image can be when partially occluded. The ratio utilised when three objects were covering it was a maximum of 60% of the original area, thus, the emoji's total area could not be more than 60% of the image. Similarly, each emoji was not allowed to overlap other emojis by more than 20% of its size.

After placing the emoji onto the image, the annotation information was saved by calling the functions mentioned earlier. However, the location at which they are saved depended on the number of images already generated. In other words, the image could be saved into the Train, Validation or Test set depending on the number of images already saved. The split utilised was a 60-20-20 one, therefore 900 images were allocated to the Train Set and 300 each for the other two sets. Finally, after saving both the images and the annotation information in all three formats, the image was also saved into the Classification directory. If the image contained inset emojis, it was placed into the inset subfolder, otherwise it was saved into the noninset subfolder.

A few other evaluation functions were also defined, one to display the frequency of the type of emoji inset into an image by utilising a bar graph and the other to display the frequency of the emoji sizes within a specific range by displaying both a histogram and a kernel density plot.

Once all functions were defined, the emoji inset creator function was called four times, each time assigning a different number of emojis to be inset, starting from zero to three. For each time the function was called to inset emojis, the evaluation graphs were also calculated and displayed. The final cumulative graphs were also depicted by summing together the values of the three inset emoji dictionaries. Figure 4.1 below depicts a flowchart to explain the main steps involved in creating the dataset.

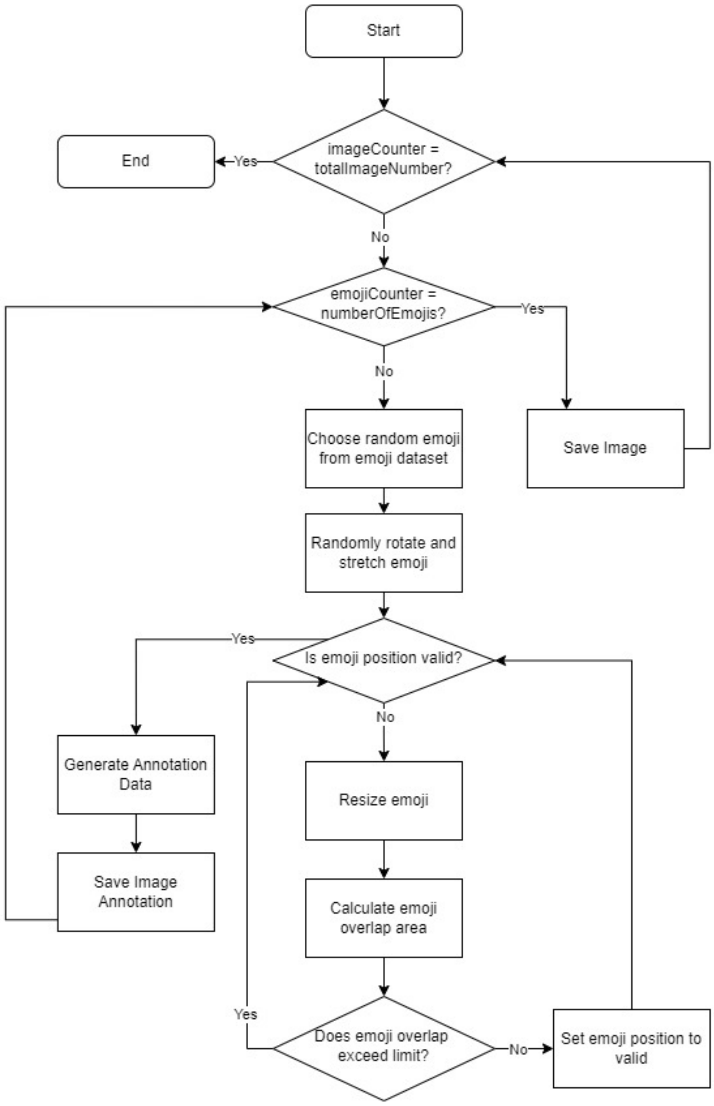


Figure 4.1 Inset Image Creator Flowchart

4.1.3 Dataset Centre Bias Evaluation

Another notebook file was utilised to evaluate the centre bias of the dataset. This notebook first imported the required libraries and updated the CSV file directories appropriately. Following this, the centre of each inset emoji was retrieved and saved within a list. Similarly, the same process was done for the image sizes in the same loop iteration. By using the centroids calculated and normalising the coordinates with respect to the image size and translating that to a new square grid, a heatmap graph was generated depicting the positions of emojis in the entire dataset. A gaussian blur effect was utilised to better understand which areas within the grid were more crowded than others. A flowchart explaining the main steps to calculate the centre bias can be seen in Figure 4.2. Also, Table 4.1 serves as an example of centre bias since it has a high frequency concentration towards the centre of the 6x6 space grid.

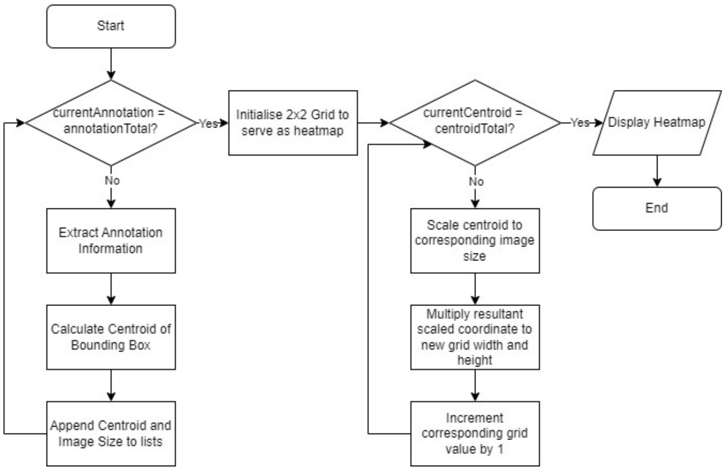


Figure 4.2 Calculate Centre Bias Flowchart

0	1	1	1	0	1
1	1	2	1	1	3
0	2	3	4	3	0
1	0	3	4	1	0
1	2	2	2	1	1
0	0	1	0	1	0

Table 4.1 Example Centre Bias Display in a 6x6 Grid Space - Each value depicts the Frequency of an Emoji Centroid being in the Coordinate

4.2 YOLOv8 Object Detection

4.2.1 Packages Required

In the YOLOv8 object detection training notebook, a few packages were required to train the model and evaluate its performance.

Object Detection

- **ultralytics.YOLO** - This module provided functionality for utilising the YOLO (You Only Look Once) object detection model.

Image Handling

- **PIL** - This module provided image processing capabilities, including loading, manipulating, and saving images.
- **IPython.display** - This module provided functions for displaying images.

4.2.2 Setting up .yaml file

For YOLOv8 object detection training, a **data.yaml** file was created, containing three paths to the respective training, validation and testing sets. The file also contains the list of label names and the number of total labels, that being seven.

4.2.3 Object Detection Notebook Walkthrough

Once the required libraries were installed, the **yolov8s.pt** model was loaded to begin training. The model was then trained on the training set, which was located thanks to the .yaml file for ten epochs. After this, the model was both validated on the validation set and tested on the testing set. The resulting graphs were automatically generated by the YOLO module in the **runs** folder.

For human evaluation, the model was given a random image from the test set and has its predictions displayed to the user, with the bounding boxes, label and confidence shown.

4.3 RetinaNet Object Detection

The RetinaNet model proved to be slightly more complex in implementation than the YOLOv8 counterpart, as more libraries and packages were required to perform the training and evaluating process.

4.3.1 Packages Required

A few modules were utilised for reasons similarly explained previously, these being **Tensorflow**, **cv2**, and **PIL**. Some other modules used were:

- **keras_retinanet.models** - This module was utilised for loading models using the RetinaNet architecture.
- **keras_retinanet.utils.visualization** - This module provided functions for visualising objects detected in images.
- **keras_retinanet.utils.image** - This module provided utility functions for preprocessing images for object detection.

4.3.2 Object Detection Notebook Walkthrough

Once the required libraries and packages were installed, a CSV file called **Classes.csv** was created based on the Training CSV file to store the label names with their associated label identifier. Following this, the **resnet50_csv_v1.h5** model was retrieved and downloaded. A few minor changes were made to the downloaded files for evaluation purposes.

The model was then trained on the training set for ten epochs, to be able to compare the results with the YOLO object detector. Following this, the model was saved, loaded and tested on the testing set. Additionally, a few examples were displayed for human evaluation.

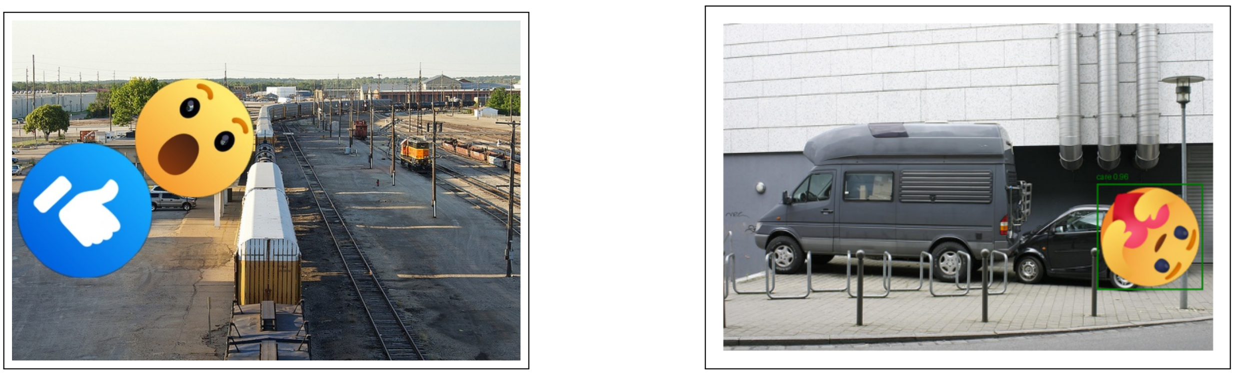
For evaluation, a mean average precision graph was depicted, each column representing the label. Furthermore, a precision-recall graph was also created to compare the model's performance on each label to the results of YOLOv8's model performance.

4.4 Reset Dataset

In the main project folder, another notebook was provided to reset the dataset, with which the folder contents could be cleared, including the images and annotations. The CSV files, rather than being deleted, simply had their contents cleared. This notebook should be used when the user wishes to create another dataset, by first clearing the dataset folder contents to be able to generate new images and annotations.

4.5 Implementation Results Examples

Once the dataset was created, the object detection models could be trained as well. The results of the implementation process were then tested in Chapter 5, which covers the evaluation aspect of this project. Figure 4.3 below depicts two examples of the resultant implementation, Figure 4.3a showing a dataset image and Figure 4.3b depicting the emojis detected by the trained RetinaNet model.



(a) Example Image from Created Dataset

(b) Emoji Detection using RetinaNet

Figure 4.3 Implementation Example Results

5 Evaluation

This chapter covers the results generated for the dataset, the object detection models trained as well as human evaluation. Each subsection explains the results and the reasoning behind why such results were generated.

5.1 Dataset Evaluation

To evaluate the dataset, three graphs were utilised, the first of which being the Emoji Type Frequency Bar Graph. This graph depicts each emoji type on the x-axis against their frequency on the y-axis, meaning the number of times the emoji type was utilised within the dataset. Due to this dataset consisting of seven emojis, seven bars are displayed in the graph. When creating the dataset, emojis were inset into images either once, twice or three times. For each of these scenarios, the graph was depicted with the data collected of that set, as well as one final time by summing the results of the sets together. The graphs generated for each set can be seen in Appendix A.2, while the finalised result can be seen in Figure 5.1.

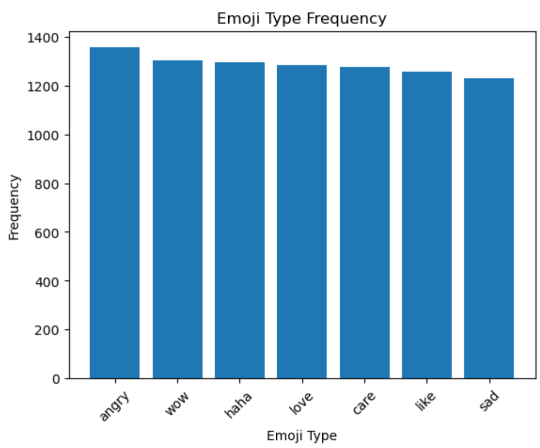


Figure 5.1 Emoji Type Frequency Bar Graph - An even distribution was created amongst all seven labels.

In all separate set figures, all labels shared a very similar frequency, with none differing from another by more than a hundred. In the final graph in Figure 5.1, the most popular emoji was the Angry one and the least was the Sad one. However, these graphs were only generated within a singular run of the dataset. If the dataset was to be created again, these graphs could have easily been different since the nature of the emoji selection is completely random.

Similar to the frequency bar graph, an emoji size distribution graph was created to evaluate the sizes of each inset emoji. A total of twenty bins were selected as it provided a general idea of the sizes of the emojis without being overly specific. The graph was also converted into a kernel density estimation plot to visualise the normal distribution generated. The resulting final graphs can be seen in Figures 5.2, where the majority of emojis tended to have a total area of approximately 0.035. The area was calculated in relation to the total area of the image, hence a total area of 0.1 would mean that the emoji occupies 10% of the total area of the image. Also, the graphs for the sets separated by the number of emojis can be found in Appendices A.3 and A.4.

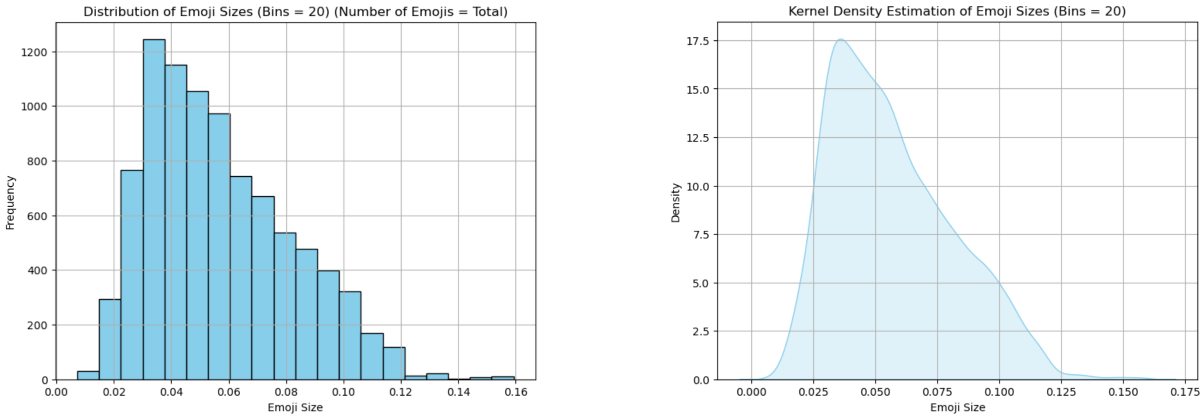


Figure 5.2 Emoji Size Histogram and Kernel Distribution Plot - A normal distribution was created.

Finally, another graph utilised to evaluate the dataset's quality was a centre bias heat-map. The reason why a center bias heat-map is used in datasets is to visualise the distribution of data points relative to a central point. This type of heat-map highlights clusters or patterns of data in relation to the image size, in this case the emoji centres. Thus, this enables the user to identify any heavy concentrations of distribution in data in a particular region of the image. The heat-map generated can be seen in Figure 5.3, which depicts a very symmetrical heat-map on all sides with no heavy concentration on the centre of the graph, a common issue with other datasets.

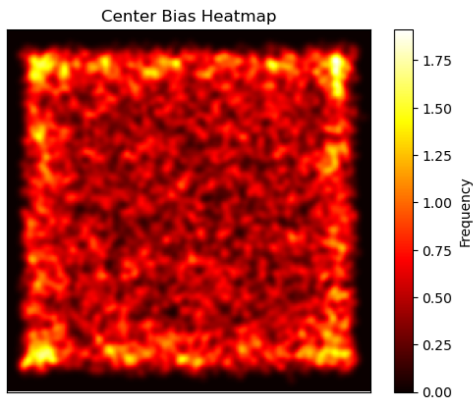


Figure 5.3 Dataset Centre Bias Graph over whole Dataset - Each point on the graph depicts the centre point of an emoji mapped to the area of this graph in relation to its original image.

5.2 Object Detectors

With regards to object detectors, YOLOv8 provided multiple different types of graphs detailing the performance of the model on both the training and test set. For evaluation purposes, the testing set was mainly utilised to analyse the model's performance due to the set consisting of unseen examples. After each run, a Recall, Precision, Precision-Recall and F1-Curve graph was saved. The resulting graphs can be seen in Figure 5.4, 5.5, 5.6 and 5.7. The model obtained very positive results in all graphs, as can even be seen in the Confusion Matrix in Figure 5.8, which shows the performance of the model in terms of True Positives, where the model achieved very high results for all labels. Apart from graphs, the model also took a sample of images from the test set and created a 4x4 image consisting of object detection results on each image, with the respective bounding boxes, labels and confidence. Three examples of these image samples can be seen in Appendix A.5.

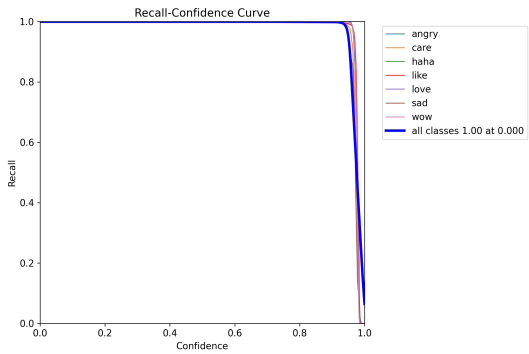


Figure 5.4 YOLO Recall Curve

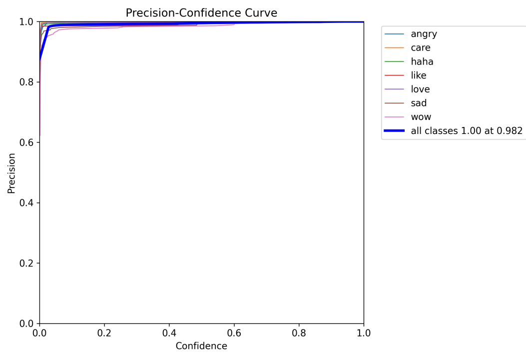


Figure 5.5 YOLO Precision Curve

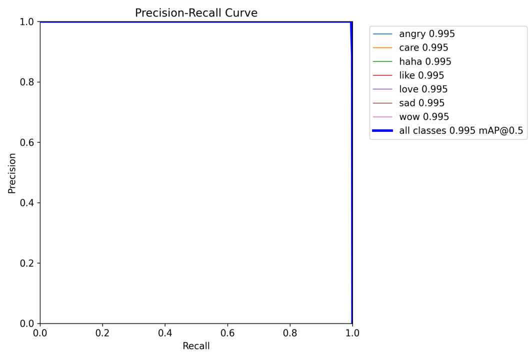


Figure 5.6 YOLO Precision-Recall Curve

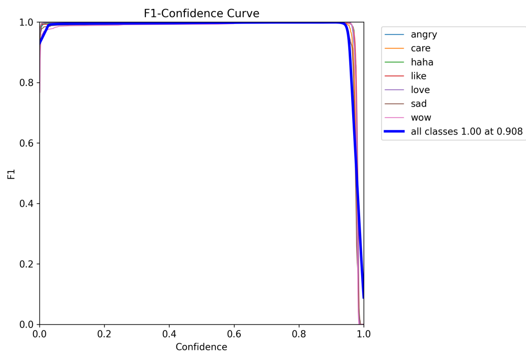


Figure 5.7 YOLO F1 Curve

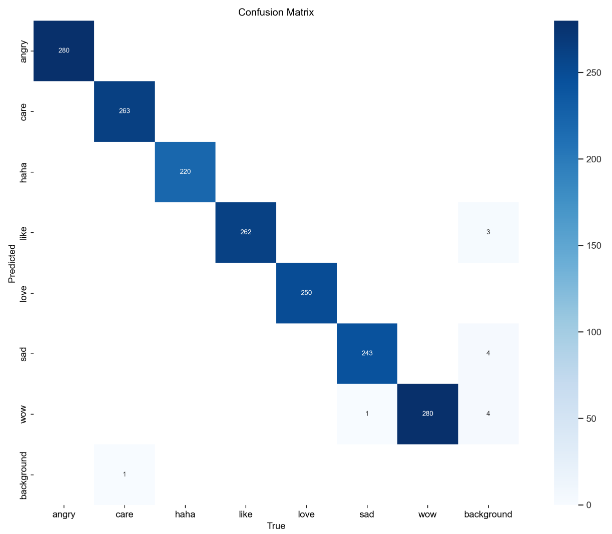


Figure 5.8 YOLO Confusion Matrix

For RetinaNet, a Precision-Recall graph was also generated to be able to compare the results to the YOLOv8 model. As can be seen in Figure 5.9 and Table 5.1, whilst RetinaNet did obtain good results, it was still not as accurate as YOLOv8's model despite being given the same training set and number of epochs. Furthermore, a Mean Average Precision bar graph is also depicted to evaluate which labels the model was best at detecting. As can be seen in Figure 5.10, the model performed best with the Love emoji, but struggled with the Haha emoji.

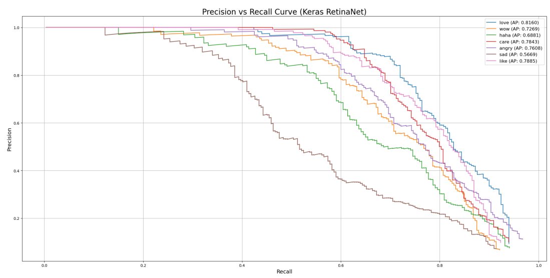


Figure 5.9 Retinanet Precision-Recall Curve

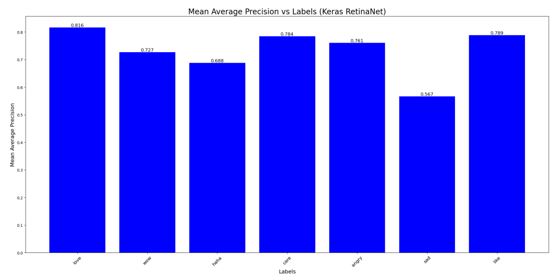


Figure 5.10 Retinanet Mean Average Precision Graph

Table 5.1 A comparison between the Average Precision Values of YOLOv8 and RetinaNet, depicting the superior results achieved by the former model.

	Love	**Wow**	**Haha**	**Care**	**Angry**	**Sad**	**Like**
YOLOv8	0.995	0.995	0.995	0.995	0.995	0.995	0.995
RetinaNet	0.816	0.727	0.688	0.784	0.761	0.567	0.789

5.3 Human Evaluation

For human evaluation, the main hypothesis was to check whether the addition of inset images has an effect on the reader. To do such a thing two types of research were carried out:

- A quantitative survey was used to examine both the general public's knowledge and awareness about the use of inset emojis in news articles.
- A qualitative interview focused on uncovering insights on the use of inset emojis in news articles from a professional in the field.

5.3.1 Survey

The survey was handed out to people at random to obtain a diverse range of different respondents from different backgrounds. This consisted of general demographic questions followed by a few questions regarding the knowledge and opinions about inset emojis in news articles. Furthermore, ten images retrieved from the dataset created were shown with and without inset emojis and the respondent was asked to select how their opinion or feelings changed.

Demographics

The survey was filled in by a total of 170 respondents, with a majority of 57.1% them being female and the other 42.9% being male as seen in Appendix A.7. In terms of age ranges, this survey was given to people of age 18 or older, and as can be seen in Appendix A.6, all age ranges were represented with the majority being aged between 18 to 26, shortly followed by those aged 43 to 51. Finally, this survey was mainly given to people of Maltese nationality, however a few Portuguese and French also filled in the survey as seen in Appendix A.8.

General Questions

From the results gathered, the majority of respondents do read news articles, with the norm being to read them sometimes or quite often. 55.3% of these readers have noticed the utilisation of inset emojis in news articles, with 50.4% of them thinking that they do not necessarily help understand a news article better. However, this does not reflect the importance of images in general, which 40.6% of respondents believe to at least be slightly important and another 20% to be very important. Despite being widely adopted and utilised, the majority of respondents believe that inset emojis detract the credibility of news articles and are inappropriate given the serious nature of such articles. Furthermore, inset emojis do not necessarily reach the audience more since most respondents do not necessarily prefer articles with inset emojis over those without and are not more likely to read an article or engage with it over those without inset emojis. All the graphs generated can be found in Appendix A.3.2.

Inset Emojis Results

This section of evaluation was utilised to undermine the changes in opinion the inset emoji can have on a reader's overall interpretation of an image. The results of this evaluation can be seen in bar chart form in Appendix A.3.3. Interestingly, for the first example shown with a Like emoji inset into the image, its presence seemed to weaken all feelings, even in terms of how calming it is. In the second image, a sharp decline in feeling happy is seen with the presence of a Cry emoji, also coupled with an increase in feeling sad. The decrease can also be seen in feeling excited and calm. The third image inset an Angry emoji, notably causing a shift in feeling more angry and less happy and calm. The fourth image consisted of a plate of food with a Love emoji inset onto it, which with and without had little to no change in feelings in all five categories. The fifth image, containing a Wow emoji seemed to only slightly alter the excited feeling to be more excited and reduced the calm feeling. The sixth image contained a Care emoji, which increased happiness whilst reducing sadness.

The seventh and eight images consisted of two emojis inset into the images. For the seventh image, a Like and Wow emoji were inset into the image, causing little change except for a slight increase in excitement. The eight image however elicited a much more noticeable change in opinion since it consisted of two Angry emojis, which made respondents feel more angry, excited and sad while decreasing their feelings of happiness and calmness they had prior to the emojis being inset.

The final two images consisted of three emojis inset into the images. The ninth image has a Cry, Care and Wow emoji inset, which made respondents less calm and happy and slightly more angry and sad. The final image inset three Care emojis onto an image of two giraffes, which surprisingly caused respondents to feel more angry and less happy.

From the results above, it can be concluded that despite emojis sometimes having the effect of altering a reader's opinion on the original image as can be seen in Image 2's results, an excessive number of emojis can sometimes cause frustration to the user as seen in Image 10's results. Furthermore, some emojis have more of an effect than others, with the Angry Emoji causing the most noticeable shifts in behaviour from the set, whereas the Like and Love emojis tend to have little to not effect. Moreover, whilst the emojis might have an impact on the image, another key factor in swaying a reader's image is the background image itself. If the background image contains something which can be interpreted vastly different with an emoji, for instance a human being, the respondents are more likely to have a greater shift in opinion with the presence of an emoji.

5.3.2 Interview

The interview was conducted with a professional in the media industry to provide insights into the selection and reasoning behind the use of inset emojis. They were asked about their experience with inset emojis and the effects that they have noted on the readers when using them. Additionally, they were also inquired on what they believe to be the future of inset emojis, whether they are just a trend that will die down or something that will become even more prevalent in the future.

Introduction

An interview was conducted with Neil Camilleri, a professional in the media industry who has been a journalist for the past 20 years and worked with multiple newsrooms. The fully transcribed interview can be found signed in Appendix B. The interview mainly focused around the use of images in news articles, particularly inset emojis, their use and their credibility within news articles.

Experience with Inset Emojis

Mr Camilleri noted that around two years ago, a few local newsrooms such as LovinMalta and MaltaDaily started utilising inset emojis, which sparked a trend. This trend led to an influx of social media-friendly, sensational posts over in-depth reporting. Such newsrooms rely on social media platforms including Instagram and Facebook, through which they overlay emojis on images in their posts to boost engagement. Some traditional outlets like Net News also experimented with this on TikTok. However, Mr Camilleri noted that the trend has waned in the past year, with platforms like LovinMalta scaling back or completely dropping inset emojis, possibly due to concerns with their effectiveness. He also added that he believes inset emojis were simply a passing trend, which can be seen on the international level too.

Whilst agreeing with the use of inset images to bring more context to the article, Mr Camilleri criticised the trivialisation of news through emojis, deeming it unprofessional and potentially manipulative. He also argued that journalism should focus on storytelling without the need of emoji influence. Even though emojis are widely utilised in daily life conversation, he opposed their integration into news, especially in serious or sensitive topics. Mr Camilleri added that the prevalence of inset emojis is mainly seen in lifestyle-oriented platforms, urging news outlets to make a clear distinction between serious journalism and entertainment-focused content, to preserve credibility.

Emoji Selection

With respect to image selection, Mr Camilleri stated that the online editor typically oversees the choice of imagery used and consults it with the journalists, aiming for impactful yet ethical visuals. The process for emoji selection is similar, normally targeting a younger demographic. However, Mr Camilleri viewed these emojis as trivial, frequently utilised in non-serious news, akin to clickbait. He advocated for a shift of focus on quality content and imagery rather than settling for inset emojis as substitutes for skilled photography. The importance of headlines and image quality are emphasised for better journalistic integrity and visual storytelling. Mr Camilleri warned against the use of inset emojis due to it compromising the article's neutrality, coming at the risk of bias. He also concluded that inset emojis further obscure the truth in reporting.

Purpose of Images and Emojis

When asked about the purpose of images, Mr Camilleri stated that visuals have a profound impact on a reader's opinion about a news article as a whole, as they give the initial impression of the article alongside the headline. He also cautioned against the possible negative consequences of inset emojis, as the initial impression created by an emoji can persist even after reading an article's content that contradicts it. Moreover, Mr Camilleri stressed the importance of journalistic restraint towards the use of emojis, as it may influence a reader's opinion. Finally, he also questioned the necessity of including these emojis, and if they truly are effective in engaging with younger audiences to foster interest in important issues such as corruption.

Audience Reactions

In terms of audience reactions, Mr Camilleri noted that emojis and clickbait images can generate quick but shallow engagement, where people react to a news article without actually reading its content. Because of this, he argued that the diversion of attention caused by inset emojis undermine the importance of the article's subject matter. Hence, a newsroom may prioritise clicks and views over the actual quality of its content due to its competitive nature of the advertising industry. While efforts to enhance visuals may increase engagement, they also come at the expense of content quality, potentially declining the overall journalistic standards.

Challenges

With regards to challenges faced by newsrooms, Mr Camilleri pointed out that many newsrooms face a problem in sourcing and selecting images for articles. A vast amount of newsrooms lack a full-time photographer, thus relying on stocks images that may not reflect the article accurately. Licensing and copyright issues only further complicate the use of images, this is why Mr Camilleri believed that local stock images are often augmented with emojis for more appeal. Another challenge mentioned is where to draw the lines on what to show in an image an what not to. He stressed that image selection is a critical aspect of article presentation, particularly when dealing with sensitive topics such as human tragedy.

Future of Inset Emojis

In terms of the future of inset emojis, Mr Camilleri has observed a decline in their overall usage over the past few months, however this is not to say that they are not popular at all. He also stated that if emojis do fade away as a trend, a new trend will replace it due to a copycat culture currently present in the media industry. This does not necessarily have to be in the visual aspect but even in literary styles, subject matter or even campaign initiatives. Finally, Mr Camilleri also discussed the emergence of Tiktok, which many newsrooms are currently experimenting with to reach younger audiences, but expresses skepticism about its effectiveness for serious news delivery. He argued that such platforms are mainly utilised for entertainment rather than news consumption. Mr Camilleri concluded the interview by criticising the current lack of creativity in journalism, where newsrooms focus on covering the same stories and compete over minor updates on a story. He attributes this behaviour to a journalist's ego and the competitive nature of the industry. Whilst some trends may lead to increased coverage of important events, he warned against blindly following trends such as inset emojis which compromise the credibility of the media sector. He emphasised the need for originality and cautioned against sacrificing journalistic integrity for the sake of following trends.

6 Conclusion

The main hypothesis of this project, that being that inset emojis have an influence on a news article reader, was proven to be true by the research results. However, it should be noted that these results were based on a small sample of the Maltese population, and thus one might question the generalisability of such a finding on an international context. In spite of this, it should be noted that the images utilised are Facebook emojis, a type of emojis available worldwide. Given this, it can be assumed that if a broader sample of the population was asked to fill in the survey, similar results would be replicated. Since the hypothesis was proven, the dataset as well as the object detection models are available to tackle media bias issues. The performance results found for both the quality of the dataset as well as the precision of the trained models proves that despite still requiring future work and research, the field of inset emojis shows great potential.

This project has provided research in two separate fields, the computer science field, specifically in the Artificial Intelligence sector, and the social science field. In the computer science field, the two main objectives related to code have been achieved. An inset emoji dataset has been provided and clearly documented to aid future projects in training models. In addition to this, two object detection models have been provided to detect inset emojis by training them on the custom dataset. This serves as both a test of the quality of the dataset as well as the models trained too. In the social science field, the human evaluation objective has also been achieved since this study has shown that emojis certainly have an effect on a viewer's interpretation of an image when inset onto them. However, this is not necessarily entirely dependent on the emoji itself, rather the type of emoji it is and how it relates to the background image.

Whilst this project provides a foundation for future work related to inset emoji detection and media bias prevention, some upgrades to the current project could be made. Firstly, with respect to the dataset itself, if more studies are made with regards to other emoji datasets, the inset emoji dataset could be created on other types of emoji datasets. Furthermore, since technology, especially Artificial Intelligence, is rapidly evolving, newly released object detection models could be trained for more accurate emoji detection once enough literature is made to back them up. Finally, for human evaluation, a qualitative approach could be utilised, however this would require much more time both to fill in and to analyse. Additionally, more interviews could be made with professionals in the media industry field to provide any insights which were not found within this study.

References

[1] *Image definition & meaning*, Merriam-Webster, Accessed: 03 May 2024. [Online]. Available: https://www.merriam-webster.com/dictionary/image.

[2] Q. Bai, Q. Dan, Z. Mu, and M. Yang, "A systematic review of emoji: Current research and future perspectives," *Frontiers in Psychology*, vol. 10, Oct. 2019. doi: 10.3389/fpsyg.2019.02221.

[3] G. Hili, *Using machine learning to investigate potential image bias in news articles*, Faculty of ICT, University of Malta, May 2023.

[4] S. Mullainathan and A. Shleifer, "Media bias," National Bureau of Economic Research, 1050 Massachusetts Avenue, Cambridge, MA 02138, Working Paper 9295, Oct. 2002.

[5] H.-W. Hu, Y.-K. Tsay, and C.-Y. Peng, "Influence of online news features on user behavior for new media," in *2018 IEEE International Conference on Big Data (Big Data)*, 2018, pp. 3659–3664. doi: 10.1109/BigData.2018.8621963.

[6] S. L. Sargent, "Image effects on selective exposure to computer-mediated news stories," *Computers in Human Behavior*, vol. 23, no. 1, pp. 705–726, Jan. 2007. doi: 10.1016/j.chb.2004.11.005.

[7] T. G. L. A. van der Meer, A. C. Kroon, and R. Vliegenthart, "Do News Media Kill? How a Biased News Reality can Overshadow Real Societal Risks, The Case of Aviation and Road Traffic Accidents," *Social Forces*, vol. 101, no. 1, pp. 506–530, Sep. 2021, issn: 0037-7732. doi: 10.1093/sf/soab114. [Online]. Available: https://doi.org/10.1093/sf/soab114.

[8] F. Hamborg, K. Donnay, and B. Gipp, "Automated identification of media bias in news articles: An interdisciplinary literature review," *International Journal on Digital Libraries*, vol. 20, no. 4, pp. 391–415, 2019. doi: 10.1007/s00799-018-0261-y. [Online]. Available: https://ejournals.um.edu.mt/login?url=https://www.proquest.com/scholarly-journals/automated-identification-media-bias-news-articles/docview/2134087503/se-2.

[9] F.-J. Rodrigo-Ginés, J. Carrillo-de-Albornoz, and L. Plaza, "A systematic review on media bias detection: What is media bias, how it is expressed, and how to detect it," *Expert Systems with Applications*, vol. 237, p. 121 641, 2024, issn: 0957-4174. doi: https://doi.org/10.1016/j.eswa.2023.121641. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417423021437.

[10] J. Balzan, *Exploring the automatic detection of bias and stance in media houses on controversial topics*, Faculty of ICT, University of Malta, 2022. [Online]. Available: `https://www.um.edu.mt/library/oar/handle/123456789/108245`.

[11] J. Zhou, C.-M. Pun, and Y. Tong, "News image steganography: A novel architecture facilitates the fake news identification," in *2020 IEEE International Conference on Visual Communications and Image Processing (VCIP)*, 2020, pp. 235–238. doi: `10.1109/VCIP49819.2020.9301846`.

[12] *Cricketing: Rubbing your feet together to fall asleep*, Dec. 2023. [Online]. Available: `https://maltadaily.mt/cricketing-rubbing-your-feet-together-to-fall-asleep/`.

[13] M. Grima, "Watch: Understanding malta's cannabis laws here are the real facts you need to know," *Lovin Malta*, 2024, Accessed: 06 May 2024. [Online]. Available: `https://lovinmalta.com/sponsored/watch-understanding-maltas-cannabis-laws-here-are-the-real-facts-you-need-to-know/`.

[14] S. Treanor and V. Nunis, "The little-known human stories behind emoji designs," *BBC News*, 2021, Accessed: 15 May 2024. [Online]. Available: `https://www.bbc.com/news/technology-58180556`.

[15] R. Puglisi and J. Snyder, "Empirical studies of media bias," in Dec. 2015, pp. 647–667, isbn: 978-0-444-62721-6. doi: `10.1016/B978-0-444-63685-0.00015-2`.

[16] M. Fanfani, M. Iuliani, F. Bellavia, C. Colombo, and A. Piva, "A vision-based fully automated approach to robust image cropping detection," *Signal Processing Image Communication*, vol. 80, Sep. 2019. doi: `10.1016/j.image.2019.115629`.

[17] M. Gentzkow, J. M. Shapiro, and D. F. Stone, "Media bias in the marketplace: Theory," *NBER Working Paper Series*, vol. 19880, 2014. doi: `10.3386/w19880`.

[18] Times of Malta. "Valletta friars call out shops and bars for excessive dirt and noise." Accessed: 20 February 2024. (2023), [Online]. Available: `https://timesofmalta.com/articles/view/valletta-friars-call-shops-bars-excessive-dirt-noise.1037393`.

[19] Times of Malta. "Trituni, not tritoni." Accessed: 20 February 2024. (2022), [Online]. Available: `https://timesofmalta.com/articles/view/Trituni-not-Tritoni.670991`.

[20] T.-Y. Lin *et al.*, *Microsoft coco: Common objects in context*, 2015. arXiv: `1405.0312 [cs.CV]`.

[21] IconScout. “Download facebook reaction icon pack.” Accessed: 20 February 2024. (), [Online]. Available: `https://iconscout.com/free-icon-pack/facebook-reactions-emoticons`.

[22] E. Grossi and M. Buscema, “Introduction to artificial neural networks,” *European journal of gastroenterology & hepatology*, vol. 19, pp. 1046–54, Jan. 2008. doi: `10.1097/MEG.0b013e3282f198a0`.

[23] S. Schmidgall *et al.*, *Brain-inspired learning in artificial neural networks: A review*, 2023. arXiv: `2305.11252 [cs.NE]`.

[24] K. O’Shea and R. Nash, *An introduction to convolutional neural networks*, 2015. arXiv: `1511.08458 [cs.NE]`.

[25] T. Liu, S. Fang, Y. Zhao, P. Wang, and J. Zhang, “Implementation of training convolutional neural networks,” Jun. 2015.

[26] K. Oksuz, B. C. Cam, S. Kalkan, and E. Akbas, “Imbalance problems in object detection: A review,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 10, 2021. doi: `10.1109/TPAMI.2020.2981890`.

[27] Z.-Q. Zhao, P. Zheng, S.-t. Xu, and X. Wu, *Object detection with deep learning: A review*, 2019. arXiv: `1807.05511 [cs.CV]`.

[28] Ò. Lorente, I. Riera, and A. Rana, *Image classification with classic and deep learning techniques*, 2021. arXiv: `2105.04895 [cs.CV]`.

[29] D. N. Thakur, P. Nagrath, R. Jain, D. Saini, N. Sharma, and J. D, *Object detection in deep surveillance*, Sep. 2021. doi: `10.21203/rs.3.rs-901583/v1`.

[30] K. Amer, M. Samy, M. Shaker, and M. ElHelw, *Deep convolutional neural network-based autonomous drone navigation*, 2019. arXiv: `1905.01657 [cs.CV]`.

[31] Z. Oufqir, A. el Abderrahmani, and K. Satori, “Novel approach for augmented reality using convolutional neural networks,” *International Journal of Advances in Soft Computing and its Applications*, vol. 14, pp. 67–78, Aug. 2022. doi: `10.15849/IJASCA.220720.05`.

[32] P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features,” in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, vol. 1, 2001, pp. I–I. doi: `10.1109/CVPR.2001.990517`.

[33] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05)*, vol. 1, 2005, 886–893 vol. 1. doi: `10.1109/CVPR.2005.177`.

[34] G. Das, H. J. Wiener, and I. Kareklas, "To emoji or not to emoji? examining the influence of emoji on consumer reactions to advertising," *Journal of Business Research*, vol. 96, pp. 147–156, 2019, issn: 0148-2963. doi: https://doi.org/10.1016/j.jbusres.2018.11.007. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0148296318305599.

[35] T. Garcia-Collart, "Emoji marketing: Strengthening the consumer brand relationship and its downstream effects," English, Copyright - Database copyright ProQuest LLC; ProQuest does not claim copyright in the individual underlying works; Last updated - 2024-02-08, Ph.D. dissertation, 2020, p. 169, isbn: 9798380093415. [Online]. Available: https://ejournals.um.edu.mt/login?url=https://www.proquest.com/dissertations-theses/emoji-marketing-strengthening-consumer-brand/docview/2847613193/se-2.

[36] S.-N. Z. Zi-Ting Yang Wen-Qi Ruan and F. Deng, "The charm of emojis: The heterogeneous effects of emoji use on users' intention to share in online tourism rumor correction," *Asia Pacific Journal of Tourism Research*, vol. 28, no. 3, pp. 212–233, 2023. doi: 10.1080/10941665.2023.2217956.

[37] K.-Y. Wang, W.-H. Chih, and A. Honora, "How the emoji use in apology messages influences customers' responses in online service recoveries: The moderating role of communication style," *International Journal of Information Management*, vol. 69, p. 102 618, 2023, issn: 0268-4012. doi: https://doi.org/10.1016/j.ijinfomgt.2022.102618. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0268401222001529.

[38] M. A. Riordan and E. Glikson, "On the hazards of the technology age: How using emojis affects perceptions of leaders," *International Journal of Business Communication*, 2020. doi: 10.1177/2329488420971690. [Online]. Available: https://doi-org.ejournals.um.edu.mt/10.1177/2329488420971690.

[39] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998. doi: 10.1109/5.726791.

[40] M. S. Hossain and K. Andersson, "Hand-drawn emoji recognition using convolutional neural network," eng, Piscataway: The Institute of Electrical and Electronics Engineers, Inc. (IEEE), 2020, p. 147.

[41] M. Bala, M. Gupta, and R. Kumar, "Emoji classification using random forest classifier," in *2022 4th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)*, 2022, pp. 728–732. doi: 10.1109/ICAC3N56670.2022.10074460.

[42] P. D. Bharati, "Emoji classification using cnn," *International Journal of Creative Research Thoughts (IJCRT)*, vol. 11, no. 4, f256–f260, Apr. 2023, issn: 2320-2882. [Online]. Available: `http://www.ijcrt.org/papers/IJCRT2304644.pdf`.

[43] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Region-based convolutional networks for accurate object detection and segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 1, pp. 142–158, 2016. doi: `10.1109/TPAMI.2015.2437384`.

[44] S. Ren, K. He, R. Girshick, and J. Sun, *Faster r-cnn: Towards real-time object detection with region proposal networks*, 2016. arXiv: `1506.01497 [cs.CV]`.

[45] W. Liu *et al.*, "SSD: single shot multibox detector," *CoRR*, vol. abs/1512.02325, 2015. arXiv: `1512.02325`. [Online]. Available: `http://arxiv.org/abs/1512.02325`.

[46] J. Redmon, S. K. Divvala, R. B. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," *CoRR*, vol. abs/1506.02640, 2015. arXiv: `1506.02640`. [Online]. Available: `http://arxiv.org/abs/1506.02640`.

[47] K. He, G. Gkioxari, P. Dollár, and R. Girshick, *Mask r-cnn*, 2018. arXiv: `1703.06870 [cs.CV]`.

[48] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, *Focal loss for dense object detection*, 2018. arXiv: `1708.02002 [cs.CV]`.

[49] M. Tan, R. Pang, and Q. V. Le, *Efficientdet: Scalable and efficient object detection*, 2020. arXiv: `1911.09070 [cs.CV]`.

[50] G. Jocher, A. Chaurasia, and J. Qiu, *Ultralytics yolov8*, version 8.0.0, 2023. [Online]. Available: `https://github.com/ultralytics/ultralytics`.

[51] S. Dey, J. Mukherjee, and S. Sural, "Stamp and logo detection from document images by finding outliers," in *2015 Fifth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG)*, 2015, pp. 1–4. doi: `10.1109/NCVPRIPG.2015.7489947`.

[52] C. Lu, D. Li, and D. Zeng, "Logo detection based on convolutional neural networks," in *2019 IEEE International Conference on Signal, Information and Data Processing (ICSIDP)*, 2019, pp. 1–6. doi: `10.1109/ICSIDP47821.2019.9173284`.

[53] K. Rezkiani, I. Nurtanio, and Syafaruddin, "Logo detection using you only look once (yolo) method," in *2022 2nd International Conference on Electronic and Electrical Engineering and Intelligent System (ICE3IS)*, 2022, pp. 29–33. doi: `10.1109/ICE3IS56585.2022.10010121`.

[54] J. Singh, K. Upreti, A. K. Gupta, N. Dave, A. Surana, and D. Mishra, “Deep learning approach for hand drawn emoji identification,” in *2022 IEEE International Conference on Current Development in Engineering and Technology (CCET)*, 2022, pp. 1–6. doi: `10.1109/CCET56606.2022.10080218`.

[55] Y. Tian, T. Galery, G. Dulcinati, E. Molimpakis, and C. Sun, “Facebook sentiment: Reactions and emojis,” in *Proceedings of the Fifth International Workshop on Natural Language Processing for Social Media*, L.-W. Ku and C.-T. Li, Eds., Valencia, Spain: Association for Computational Linguistics, Apr. 2017, pp. 11–16. doi: `10.18653/v1/W17-1102`. [Online]. Available: `https://aclanthology.org/W17-1102`.

[56] G. Das, H. J. Wiener, and I. Kareklas, “To emoji or not to emoji? examining the influence of emoji on consumer reactions to advertising,” *Journal of Business Research*, vol. 96, pp. 147–156, 2019, issn: 0148-2963. doi: `https://doi.org/10.1016/j.jbusres.2018.11.007`. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S0148296318305599`.

[57] E. Ko, D. Kim, and G. Kim, “Influence of emojis on user engagement in brand-related user generated content,” *Computers in Human Behavior*, vol. 136, p. 107 387, 2022, issn: 0747-5632. doi: `https://doi.org/10.1016/j.chb.2022.107387`. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S0747563222002096`.

[58] M. Wibowo, R. Ats-Tsiqoh, S. Sangadah, E. Komala, and A. Utomo, “The effect of emoji on person perception,” Nov. 2016. [Online]. Available: `https://www.researchgate.net/publication/317170695_The_Effect_of_Emoji_on_Person_Perception`.

[59] T. Maiberger, D. Schindler, and N. Koschate-Fischer, “Let’s face it: When and how facial emojis increase the persuasiveness of electronic word of mouth,” *Journal of the Academy of Marketing Science*, vol. 52, no. 1, pp. 119–139, Jan. 2024, issn: 1552-7824. doi: `10.1007/s11747-023-00932-8`.

[60] V. Nithyashree. “Step-by-step guide for image classification on custom datasets.” Accessed: 02 May 2024, Analytics Vidhya. (2021), [Online]. Available: `https://www.analyticsvidhya.com/blog/2021/07/step-by-step-guide-for-image-classification-on-custom-datasets/`.

[61] J. Solawetz and Francesco, *What is yolov8? the ultimate guide. [2024]*, `https://blog.roboflow.com/whats-new-in-yolov8/`, Roboflow Blog, 2023.

[62] Keras Team, *Keras Documentation: Object Detection with RetinaNet*, `https://keras.io/examples/vision/retinanet/`, Keras.

[63] Y. Niu, F. Liu, L. Xueqing, H. Bao, and M. Gleicher, “Detection of image stretching,” Jul. 2010, pp. 93–100. doi: 10.1145/1836248.1836266.

[64] J. Wang, Z. Zhang, C. Xie, J. Zhu, L. Xie, and A. Yuille, “Detecting semantic parts on partially occluded objects,” Jan. 2017. doi: 10.5244/C.31.73.

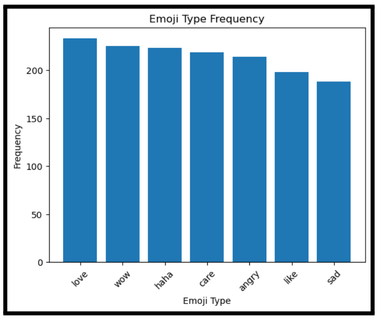
Appendix A

A.1 Annotations

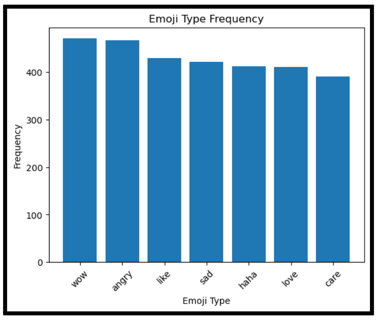
```xml
<annotation>
   <folder>EmojiDataset</folder>
   <filename>Emojis_3_Image311</filename>
   <path>labels/Emojis_3_Image311</path>
   <size>
      <width>640</width>
      <height>480</height>
      <depth>3</depth>
   </size>
   <object>
      <name>care</name>
      <bndbox>
         <xmin>377</xmin>
         <ymin>198</ymin>
         <xmax>496</xmax>
         <ymax>317</ymax>
      </bndbox>
   </object>
   <object>
      <name>love</name>
      <bndbox>
         <xmin>397</xmin>
         <ymin>279</ymin>
         <xmax>520</xmax>
         <ymax>402</ymax>
      </bndbox>
   </object>
   <object>
      <name>sad</name>
      <bndbox>
         <xmin>0</xmin>
         <ymin>277</ymin>
         <xmax>138</xmax>
         <ymax>435</ymax>
      </bndbox>
   </object>
</annotation>
```

Figure A.1 VOC XML Annotation Example

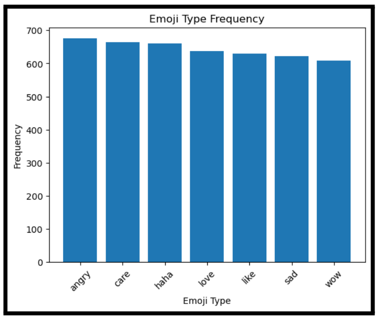
A.2 Code Evaluation



(a) Emoji Type Frequency Bar Graph (1 Emoji)

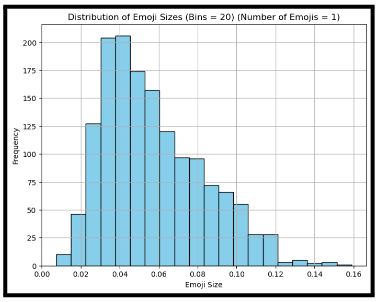


(b) Emoji Type Frequency Bar Graph (2 Emojis)

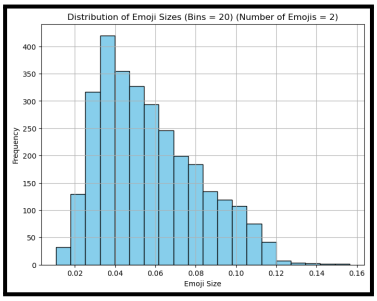


(c) Emoji Type Frequency Bar Graph (3 Emojis)

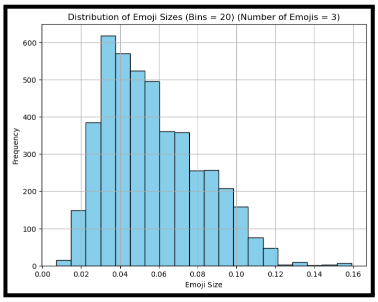
Figure A.2 Emoji Type Frequency Bar Graph Comparison



(a) Emoji Size Histogram (1 Emoji)

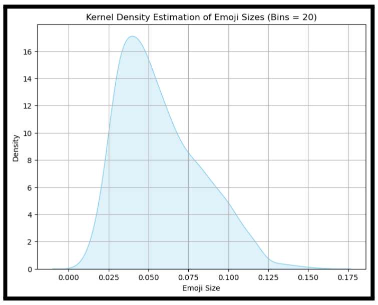


(b) Emoji Size Histogram (2 Emojis)

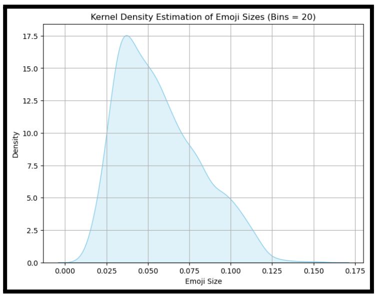


(c) Emoji Size Histogram (3 Emojis)

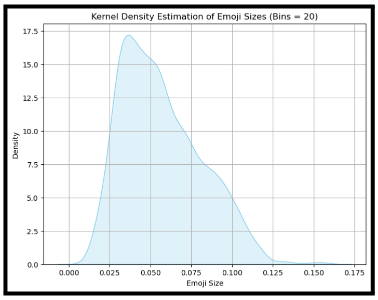
Figure A.3 Emoji Size Histogram Comparison



(a) Emoji Size Kernel Density Plot (1 Emoji)

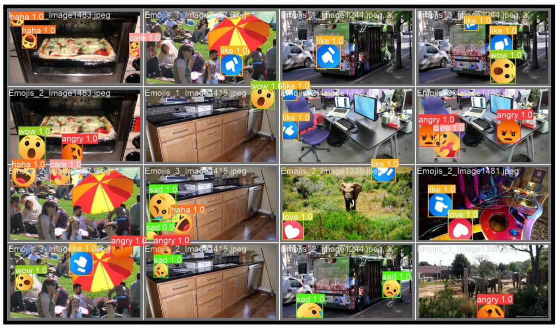


(b) Emoji Size Kernel Density Plot (2 Emojis)

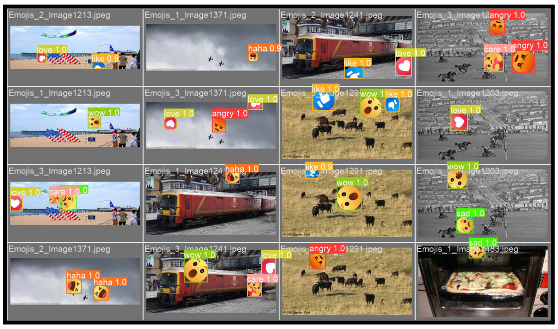


(c) Emoji Size Kernel Density Plot (3 Emojis)

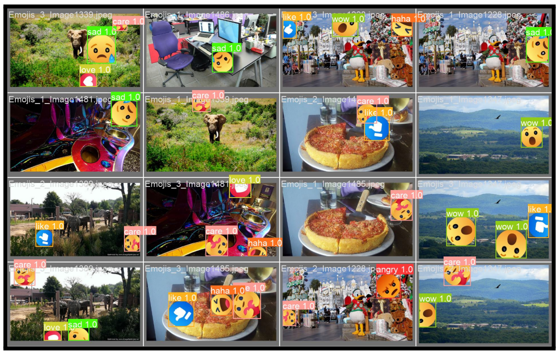
Figure A.4 Emoji Size Kernel Density Plot Comparison



(a) YOLO Sample Results (Batch 1)



(b) YOLO Sample Results (Batch 2)



(c) YOLO Sample Results (Batch 3)

Figure A.5 YOLO Sample Results

A.3 Survey Results

A.3.1 Demographics

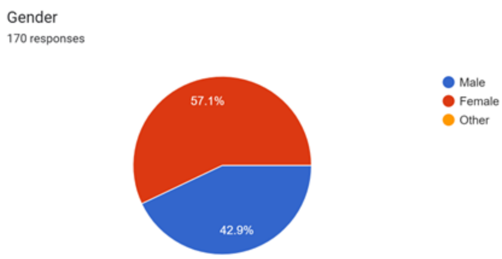


Figure A.6 Survey Results - Gender

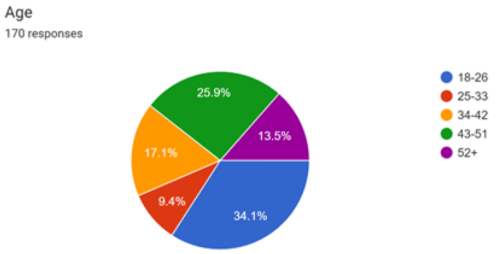


Figure A.7 Survey Results - Age

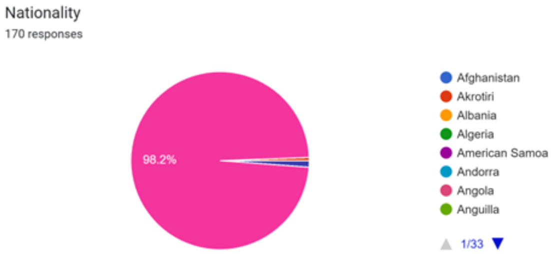


Figure A.8 Survey Results - Nationality

A.3.2 General Questions

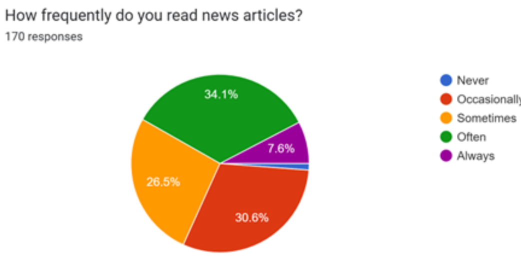


Figure A.9 Survey Results - Article Reading Frequency

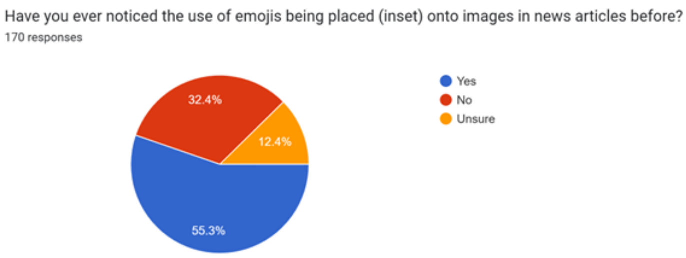


Figure A.10 Survey Results - Awareness of Inset Emojis

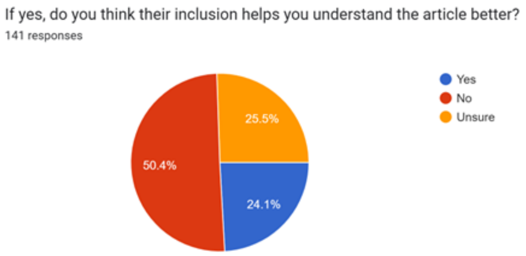


Figure A.11 Survey Results - Inset Emojis influence on Understanding

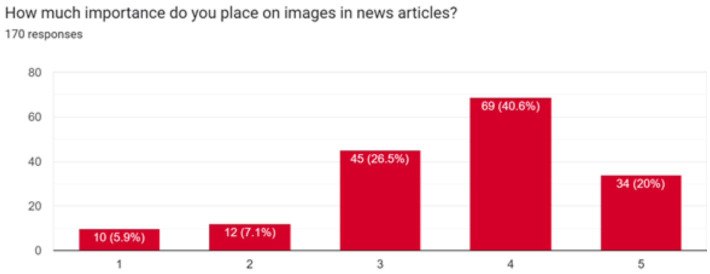


Figure A.12 Survey Results - Image Importance

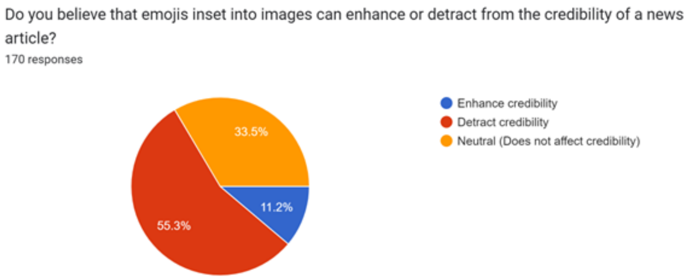


Figure A.13 Survey Results - Inset Emoji influence on Credibility

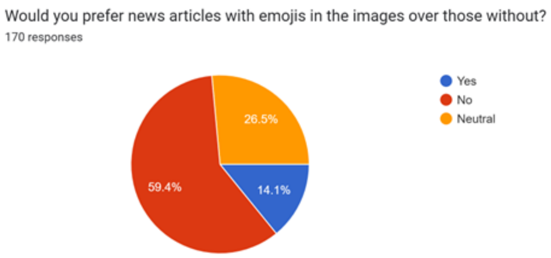


Figure A.14 Survey Results - Inset Emoji Preference

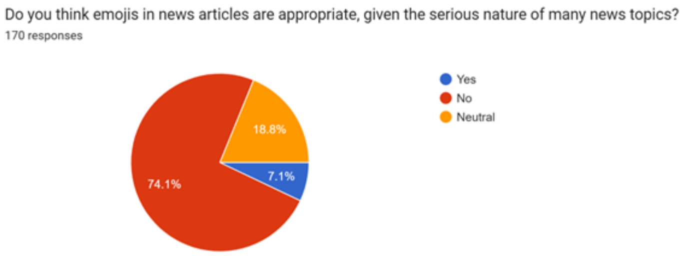


Figure A.15 Survey Results - Inset Emoji Appropriateness

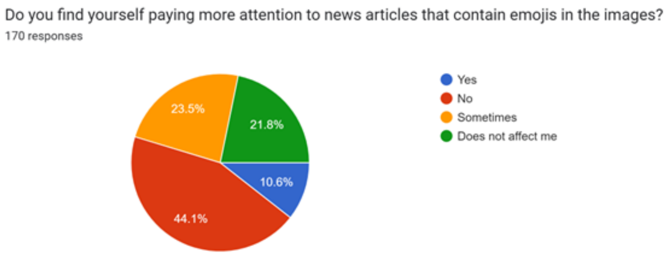


Figure A.16 Survey Results - Inset Emoji influences Attention

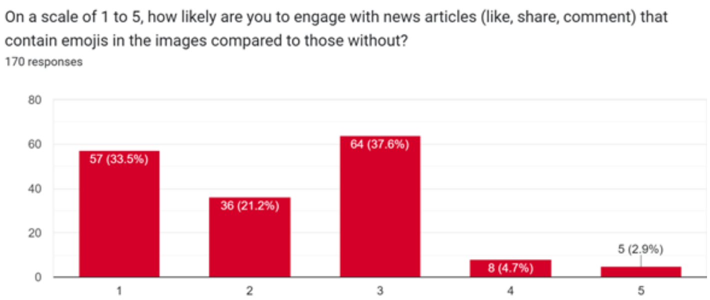


Figure A.17 Survey Results - Inset Emoji Engagement

A.3.3 Inset vs Non-Inset Comparison

Figure A.18 Survey Non-Inset Image 1

Figure A.19 Survey Inset Image 1

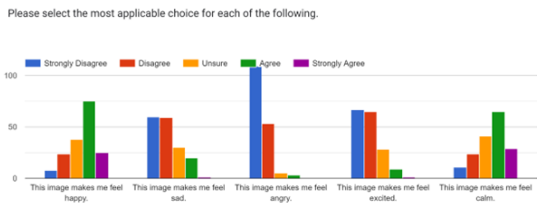


Figure A.20 Survey Non-Inset Image 1 Results

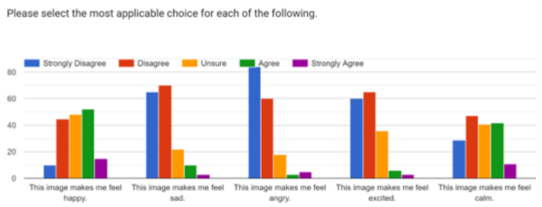


Figure A.21 Survey Inset Image 1 Results

Figure A.22 Survey Non-Inset Image 2

Figure A.23 Survey Inset Image 2

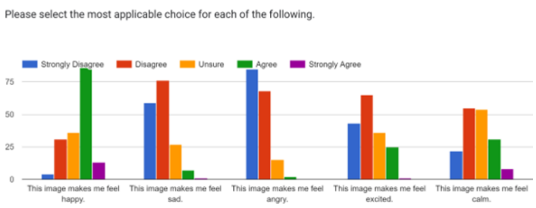


Figure A.24 Survey Non-Inset Image 2 Results

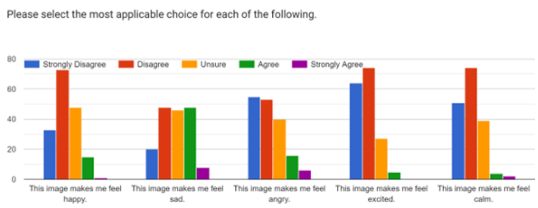


Figure A.25 Survey Inset Image 2 Results

Figure A.26 Survey Non-Inset Image 3

Figure A.27 Survey Inset Image 3

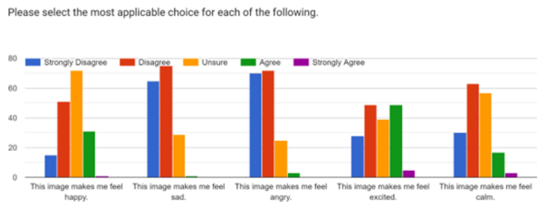


Figure A.28 Survey Non-Inset Image 3 Results

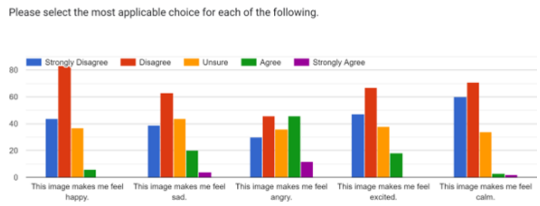


Figure A.29 Survey Inset Image 3 Results

Figure A.30 Survey Non-Inset Image 4

Figure A.31 Survey Inset Image 4

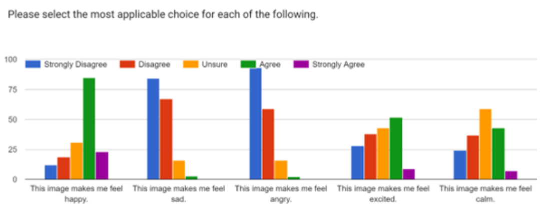


Figure A.32 Survey Non-Inset Image 4 Results

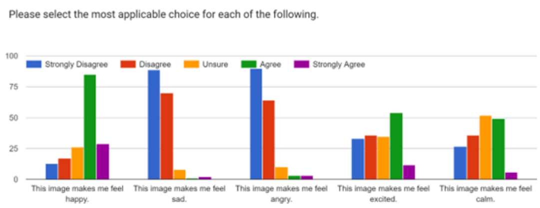


Figure A.33 Survey Inset Image 4 Results

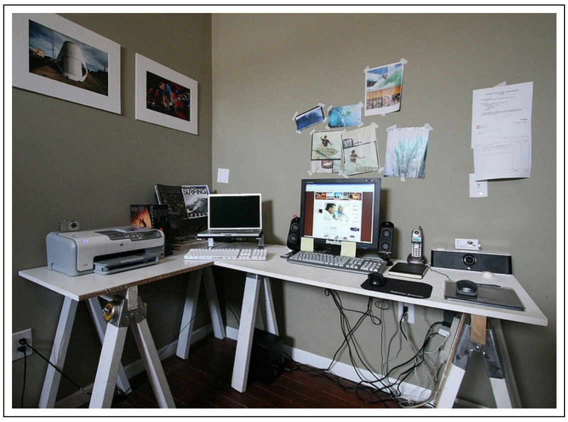


Figure A.34 Survey Non-Inset Image 5

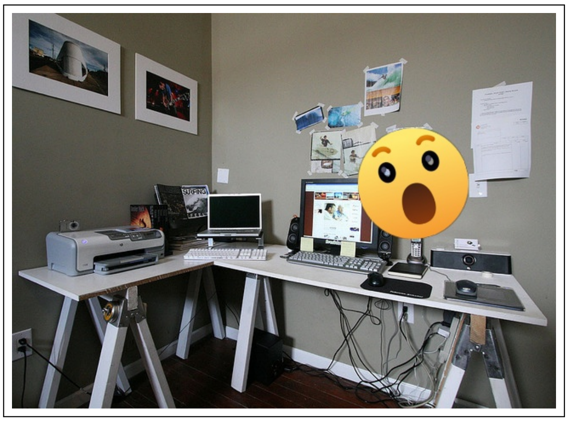


Figure A.35 Survey Inset Image 5

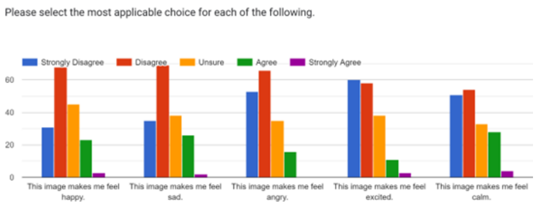


Figure A.36 Survey Non-Inset Image 5 Results

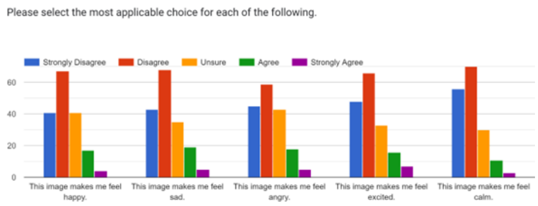


Figure A.37 Survey Inset Image 5 Results

Figure A.38 Survey Non-Inset Image 6

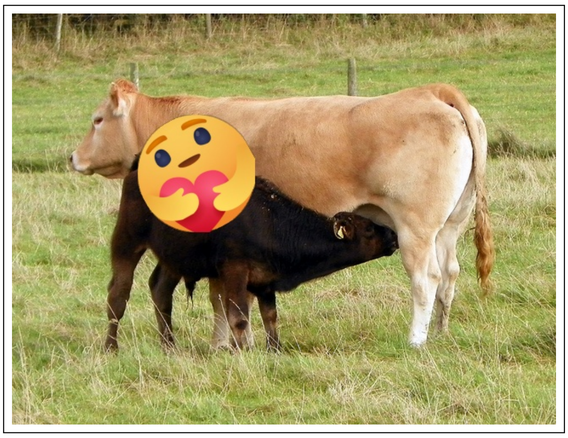


Figure A.39 Survey Inset Image 6

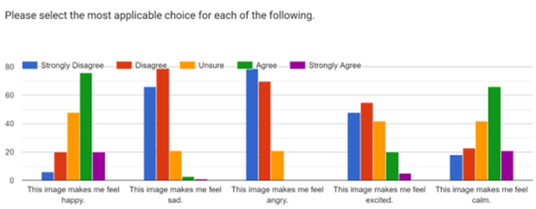


Figure A.40 Survey Non-Inset Image 6 Results

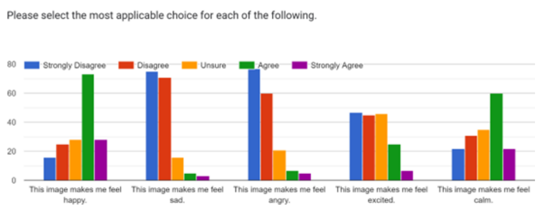


Figure A.41 Survey Inset Image 6 Results

Figure A.42 Survey Non-Inset Image 7

Figure A.43 Survey Inset Image 7

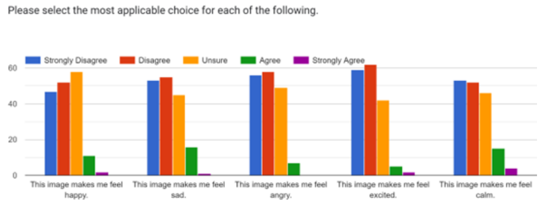


Figure A.44 Survey Non-Inset Image 7 Results

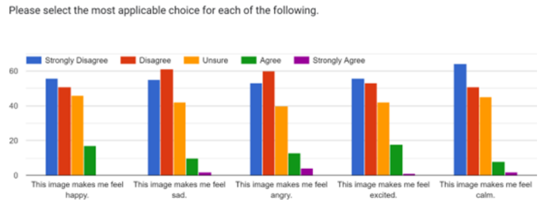


Figure A.45 Survey Inset Image 7 Results

Figure A.46 Survey Non-Inset Image 8

Figure A.47 Survey Inset Image 8

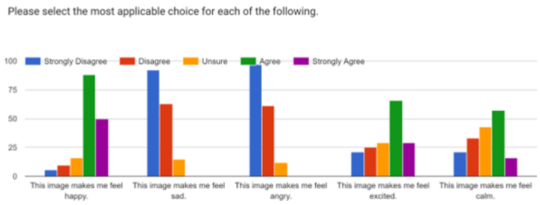


Figure A.48 Survey Non-Inset Image 8 Results

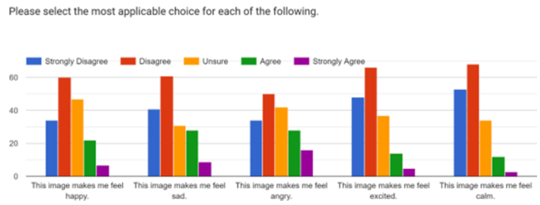


Figure A.49 Survey Inset Image 8 Results

Figure A.50 Survey Non-Inset Image 9

Figure A.51 Survey Inset Image 9

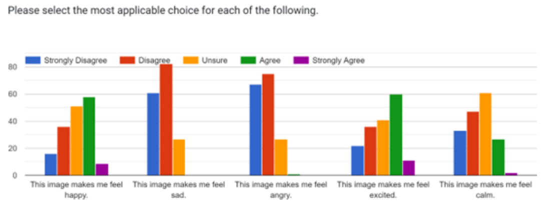


Figure A.52 Survey Non-Inset Image 9 Results

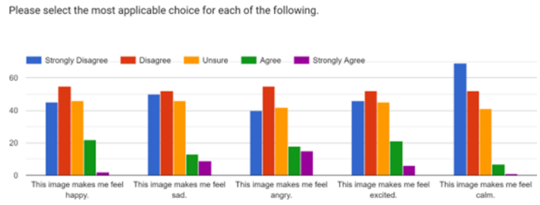


Figure A.53 Survey Inset Image 9 Results

Figure A.54 Survey Non-Inset Image 10

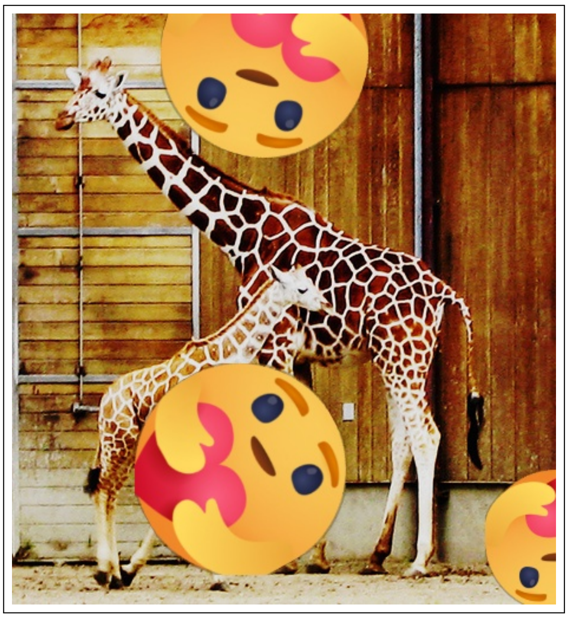


Figure A.55 Survey Inset Image 10

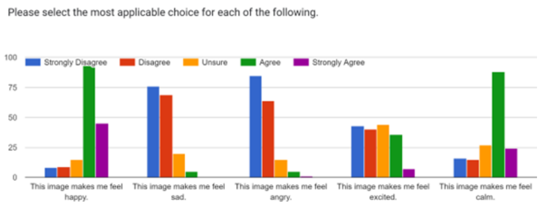


Figure A.56 Survey Non-Inset Image 10 Results

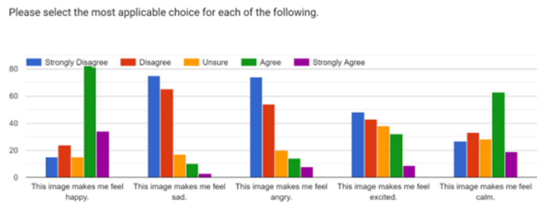


Figure A.57 Survey Inset Image 10 Results

Appendix B

B.1 Transcription

00:00:00 ISAAC MUSCAT
So hello. And with that we can start the interview. So could you please give a general introduction about yourself, your name and your role in the industry?

00:00:10 NEIL CAMILLERI
Neil Camilleri, 38 years old, journalist for almost 20 years. I've worked in radio, tv, print, online, in radio station at 89.7 Bay for two years, then Medialink communications. For six years I was a journalist, then deputy editor and a newscaster, then nine years at the Malta Independent. The last three years I was editor in chief and currently freelance journalist and media producer.

00:00:40 ISAAC MUSCAT
Okay, with that we can delve into more of the my thesis related subjects. So I would like to ask you about your experience with images in news articles, and perhaps even inset images such as emojis. So could you tell me how widely are they used in the industry over the past few years?

00:01:02 NEIL CAMILLERI
I think we had a phase probably around one or two years ago where there was a bit of a craze locally. I think this was mostly apparent on social media posts. Most of the websites or news organisations that were doing this were not strictly the traditional ones, but more of the LovinMalta, MaltaDaily kind of platforms which technically speaking, they, in my opinion, they do not fall under traditional media, but they do also present a kind of news. So we have to see them under, under the same umbrella, kind of. The difference is that they focus on very short social media posts, not like investigative or in depth reporting most of the time.

00:01:51 NEIL CAMILLERI
These two platforms, particularly LovinMalta and MaltaDaily, they're strongest on social media. Although they have their native websites, it's not their main audience base. Their main audience base is on social media, particularly, I think, on Instagram, but also to a certain extent on Facebook. And in my opinion, I mean, I have noted that these images, which are usually like traditional image related to the subject, superimposed with one or more emojis, they use them across the board on both their

websites, Instagram, Facebook and so on.

00:02:37 NEIL CAMILLERI
Having said this, I think even some of the more traditional newsrooms have used the system. One example that comes to mind is Net news, especially on social media. I don't believe they did it on their own website, but on social media there have been instances where they use these emojis, not to mention their use of TikTok to try and present news in a funny way, or in a way to try to ridicule mostly government politicians. I'm not aware if the Labour party media is the same, but yes, mostly it was these kind of lifestyle daily media platforms that do it. I have also noticed that the trend has pretty much gone away now, or decreased a lot.

00:03:36 NEIL CAMILLERI
It's still being practised specifically by Malta daily. LovinMalta seem to have stopped this practice, or if they're doing it, they're doing it very sporadically now. I'm not sure if it's because there was a change in direction, because they. I don't know, they realised that this practice doesn't make sense because in my opinion, it doesn't make. Make so much sense.

00:04:00 NEIL CAMILLERI
But even. Even on an international level, I think it was a phase that is slowly dying out now.

00:04:07 ISAAC MUSCAT
So do you think there is a difference between utilising an inset image in general or it being specifically an emoji? Do you think it's utilised for different reasons, perhaps.

00:04:21 NEIL CAMILLERI
Inset images per se I don't have a problem with. I don't have a problem of, for example, something which is very common practice of, say there was an accident, someone died, and you put an main photo of the accident site, and then as an inset picture, you do the face of the victim. That I don't have a problem with because it gives more context to the story.

00:04:42 NEIL CAMILLERI
And you eliminate this limiting factor of only using one image, because obviously the main image, the one that is going to show up on people's feeds or on people's website, on the website, is the most important one. So the more information you get inside that image, in my opinion, the better. The problem with using emojis, mainly, is that it

dumbs down the news visually. It makes it look, in my opinion, a bit childish and not professional.

00:05:17 NEIL CAMILLERI
The other problem is that it can also send out a message, because that's, at the end of the day, what emojis are. It's a comic image that conveys a certain message, usually about how you're feeling. Now, the problem is this. If I do a story and I put up an image with an emoji representing a feeling, angry, sad, surprised, whose feeling is that? Is it mine?

00:05:44 NEIL CAMILLERI
The author? Is it supposed to reflect the feeling that's going to be felt by the readers? Is it what I want them to feel? Is it what I expect them to feel? What is it exactly?

00:05:55 NEIL CAMILLERI
Who decides what this emoji should be? So I am against any sort of push by the journalist, the author, whatever, the creator, to try and feed to the audience what they should feel about the news. I think everyone has a brain and everyone can decide on their own how this news item should make them feel. If it's an accident and someone died, it's obvious. I don't need to put like a sad emoji to try and create more feeling about this news item.

00:06:36 NEIL CAMILLERI
I think the photo and the story on its own should be enough. There are other ways to engage your audience, including with good quality imagery that really reflects what's happening. Something which, for example, stock images don't really do. A lot of local news websites have a stock image base of ambulances, of hospital, of police, of court, and they always use the same ones. And that doesn't really help people connect with the story, I think.

00:07:11 NEIL CAMILLERI
But on the other hand, you shouldn't replace this with the use of emojis, especially in that kind of situation where it's like some tragic situation. I think it's quite frankly a bit offensive in the case of an accident, of a tragic accident. Obviously there are other instances, like most of local news is politics, really. But again, I don't really see a place for emojis in news language. Now, having said all this, I understand that emojis are basically a language on their own.

00:07:47 NEIL CAMILLERI

It represents a big chunk of society which are our youths. It's something we use on a daily basis. Even myself, you know, when we communicate with our peers, we tend to use these emojis a lot. It's a quick way of expressing a feeling or an emotion without going into a lot of description about it, you know.

00:08:14 NEIL CAMILLERI
So I do acknowledge that it's basically a part of our daily lives. But to put it in use, imagery, I think it's a bit unnecessary and unwarranted, to be honest.

00:08:30 ISAAC MUSCAT
So have you noticed any contexts or particular topics, perhaps where the use of said inset images or emojis is more than other topics or situations, for instance? Or is it perhaps random what is being covered?

00:08:49 NEIL CAMILLERI
I think there are different sections of news where emojis are used heavily, like I mentioned before, politics and accidents. But also, and this is connected to what I said before, that this practise is mostly used by these lifestyle daily like social media platforms, I think usually it's more connected to lifestyle news, if you want to call it that. Like these kind of spicy stories to get clicks about some celebrities or about some big events coming to Malta or things like that.

00:09:28 NEIL CAMILLERI
I think they're used more frequently in these non-serious news items than in serious ones and it tends to get a bit annoying. But this is like another pet peeve of mine is that in my opinion, you're either a newsroom or you aren't. You cannot be part newsroom, part lifestyle parts. I don't know, celebrity news and things like that. So, yeah.

00:09:57 ISAAC MUSCAT
Okay. So moving on to the choice of image or inset image selection. How would you describe the process behind finding and obtaining the image to be placed into an article, for instance.

00:10:13 NEIL CAMILLERI
Sorry, can you repeat?

00:10:15 ISAAC MUSCAT
So, with regards to the image, when we're placing an image into an article, for instance, or an inset image, what is the process or reasoning behind first obtaining it and then

choosing where to place it, for instance, within the article?

00:10:33 NEIL CAMILLERI
I wouldn't really know because I have never actually done it. I mean, what I can tell you about my personal experience with choosing images is it's been basically the prerogative most of the time of the online editor. In discussion with the journalist, you tend to choose the image which has the highest impact on the viewer, obviously within limits. For example, something which we did from the Ukraine war is we chose images that were powerful without going overboard and doing these really graphic images, which we had plenty of, but we decided not to use because obviously there's a standard, there's a code of ethics, and we don't want to cross that line and go for the shock factor. I guess it would be kind of the same lines with these emojis.

00:11:27 NEIL CAMILLERI
I think, first of all, at some point someone decided, look, let's start putting these emojis and these new stories, how they decide specifically in each image or in each story where to place them. I'm not sure about that, to be honest. I know they're going probably for maximum effect and for maximum engagement, especially with the youth readership section. I don't know. I mean, what I've noticed is that the use of these emojis is quite prominent.

00:12:00 NEIL CAMILLERI
I mean, the emojis per se are quite big, usually superimposed on the actual images. And they're there like, basically to catch your eye. Now the question is, catch your eye for what? Because if it's to get your attention, for you to read some really important article, then maybe I can kind of understand it. But like I said before, I believe most of the times they're used for more trivial stuff.

00:12:33 NEIL CAMILLERI
So having said this, they have been used, for example, with stories of a political nature. Again, I think it kind of belittles the user, the platform that is actually using it in my opinion. I might be a bit old fashioned about this. It might work a bit better with people who are younger than me. But, yeah, I mean, there's a very clear, there is a very clear intent to just get engagement solely through the image and the use of these emojis.

00:13:13 NEIL CAMILLERI
It's akin to clickbait, in my opinion, because I honestly believe that the way to get more readers is, firstly through better content, good headlines, which are not sensational, but a bit exciting and interesting at the same time, and through imagery. I think the image

per se is like one of the most important pull factors when, when we're talking about online news especially. But I think emojis do not account, do not make up for a lack of good imagery, which in my opinion can only, you can only get if you engage a good skilled photographer, basically.

00:13:57 ISAAC MUSCAT
Okay, so with regards to images in general, not specifically emojis, would you say that the selection of them differs based on the tone of the article or the demographic it's intended to and how so in what way does it usually change?

00:14:15 NEIL CAMILLERI
Yeah, definitely. It definitely reflects the tone of the article. For example, one of the most common emojis used is the angry one, or the one that is swearing. For example.

00:14:28 NEIL CAMILLERI
It's like to convey frustration, anger, disapproval. But then again, the reason why I disagree with it so much is because it basically reflects the opinion of the author, which in journalism should never happen unless it's like a signed opinion piece. You should never really give your opinion about the subject. You can kind of do it in different ways from the way you write the article. I'm not saying you should, but there are other ways of kind of leaning towards one side or the other.

00:15:11 NEIL CAMILLERI
But once you put, let's say you put an article about some speech by the prime minister and you put this angry or swearing emoji, you're basically putting your stamp on it. You know, you're telling people how they should feel when they read this, and you're also telling people how you feel, and not everyone's gonna agree with you. I think journalists already have a big problem with credibility. And to just put your signature or your stamp of approval or disapproval on your news, the news that you're presenting in such a blatant way, I don't agree with that at all.

00:15:50 ISAAC MUSCAT
Okay. And do you think this can also be reflected in just images in general without the emojis? Or is it specifically the emojis which are used to convey a certain tone of the article?

00:16:05 NEIL CAMILLERI
Look, it happens even before the advent of emojis, this was already happening to some extent. For example, I remember a time when I worked for a political party station, and

we used to be directed to choose specific photos of the prime minister at the time, photos which made him look angry, photos which made him look fat. They used to even direct the photographers to film him from, from down below to make his chin look bigger and this stupid kind of stuff to like. Even, for example, you can go to a speech, a 1 hour long speech, and use just one frame of it where the person looks a bit angry or irritated. And then you can use that photo in every single one of your articles to kind of give the impression that this guy, he's always angry, he's always looking mean, aggressive, you know, even if he was speaking about puppies.

00:17:10 NEIL CAMILLERI
So the choice of image, or rather the intentional choice of such images, has always been there, particularly in the party media. Then the use of emojis superimposed on such an image are a step further, you know? So I disagree with the first step of using these, these very selected images, but then to put an emoji on it, to make this person look even more stupid or even more arrogant or whatever, it's like doctoring to a higher degree, in my opinion.

00:17:51 ISAAC MUSCAT
Okay, so moving on to the purpose of the use of images or emojis. How would you say that the use of even images in general would influence the reader? In what way does it differ from just the standard text of the article itself?

00:18:11 NEIL CAMILLERI
I think visuals have a more profound impact on the reader in general. I think it's obviously the first thing you see, so it's going to give you the first impression about the article. I think the negative impact of it is that if you first see the image with this emoji, and you get your first impression about the article, and then you actually read the article, and you figure out that the article doesn't really reflect the emotion that the image is conveying. The image will still be stuck in your head. So it's like the first thing you see is gonna stick more.

00:18:52 NEIL CAMILLERI
So I think it's gonna precondition the reader a bit when reading the article. This is not just about, obviously, there are different, different kinds of emojis. It can make you. An emoji can make you feel sorry for someone, it can make you feel sad, it can make you feel frustrated or disgusted. You know, there's a wide range of emotions that it can make you feel.

00:19:24 NEIL CAMILLERI

Like I said, I think the first impression is what counts the most. That's why we, as journalists, need to be more restrained about the use of them. Because inadvertently, maybe, we are telling our readers to view or to digest this article, this news report, in a certain way that is maybe not the same way, you know, the way that we intend them to read it. I mean, technically speaking, what we would like our audience to do is to read it in an unbiased way and then make up their mind about it. Now, I'm not childish enough to say that the journalists cannot give some nuance of something.

00:20:14 NEIL CAMILLERI
If you're talking about corruption, you don't want people to have an open mind about it and then at the end of the day decide that it's okay, it's good. No, you want people to realise that corruption is bad, but at the same time, you cannot just like literally spoon feed these emotions to people through the use of emojis. You know, and again, there's on top of everything, the, the concept that if you're talking about such a serious subject as corruption, as government corruption, why are you using emojis in the first place? You know, it's like, do we really need to use emojis to get our young readers engaged? And are they actually gonna get engaged just because they see an emoji?

00:21:03 NEIL CAMILLERI
Are they gonna be more interested about corruption, for example, just because we use it? Because I don't really believe so.

00:21:10 ISAAC MUSCAT
So have you noticed any perhaps measurable impact on when an image or an emoji in set image is used? Or does it remain the same as without?

00:21:22 NEIL CAMILLERI
I think in some cases it gets more engagement in the sense of comments, like quick engagement, I call it, or shallow engagement in the sense that it might make more of an impact because people might rush to comment and like and share, but doesn't necessarily mean that they're digesting what the text is saying. So in a way, I feel that the use of emojis and certain, certain clickbait images can actually take the attention away from the actual text, which is at the end of the day, the most important thing. So yeah, I think it can kind of deflect attention. It creates unnecessary attention on the headline maybe, and on the picture itself, but not necessarily on the subject of the article. So it's a way of getting clicks and views, which is unfortunately what a lot of media platforms are after nowadays.

00:22:31 NEIL CAMILLERI

Because like I was saying earlier, it's like a very cutthroat industry for the advertising share. And this creates this drive to get more clicks, unfortunately, is leading to more kind of innovative ways to make the content visually more presentable, more interesting, whatever. But the content per se is not really improving. It's probably actually going down the level of content.

00:23:06 ISAAC MUSCAT
And so you're talking about assessing audience reactions, using likes and comments, for instance, is there, is that the general way that you assess the reaction to an article, or are there some other ways that you didn't mention before.

00:23:26 NEIL CAMILLERI
No, for me that's pretty much it, because I don't really, I'm not looking into or studying this subject per se, so I'm never engaged in some way to measure the impact of this. I'm glad you're doing it because someone has to do it, because I think it's a very important subject to analyse. I think it's something that not only the maltese media, but the global media, at least for this particular phase, went crazy on. But I don't think anyone has actually looked into the repercussions or actually the effects or how it's affecting, whether the important content, which is the text, is actually getting across in the same way, in an effective way. We know for sure that visually generates more interest, but I'm not sure if anyone knows whether that's translating into better news dissemination at the end of the day.

00:24:27 ISAAC MUSCAT
And so now we'll move on to the challenges that come with selecting an image. Are there any major difficulties, perhaps that are experienced when choosing to incorporate an image into an article?

00:24:43 NEIL CAMILLERI
I mean, probably the biggest problem would be a lack of base images, depending on what we're talking about. But for example, this might come as a surprise, but a lot of newsrooms don't have actually a full time photographer nowadays. For example, the newsroom I used to work with hasn't had a photographer of a geographer for a year and a half now. And this is a newsroom that produces a daily newspaper apart from the Sunday newspaper, a business newspaper and a 24/7 online presence. And it doesn't have a photographer.

00:25:23 NEIL CAMILLERI
So getting the stock based images is the biggest problem. Secondly, okay, you can use

stock images, but there is an issue of licencing and copywriting. You can just, you know, go online and steal any image that comes across. In most cases or a lot of times, to find a stock image that actually reflects the situation is going to be difficult because these stock photos are not tailor-made for you.

00:25:55 NEIL CAMILLERI
So I think what we see most of the time is a regurgitation of local stock images and then to make them a bit more enticing, some people put these emojis on them to try and make them look a bit different, a bit better. So no, I mean, look, from experience, the most difficult part about presenting an article is not writing it, it's choosing the images or finding the images. Unfortunately, there's no like ocean of photos and videos to choose from.

00:26:32 NEIL CAMILLERI
It's not only about the availability of images, it's also about choosing which one is, you know, which is the best main photo for example. Like I said before, there are some cases where you need to be careful where you might have an image which is, in your view, really, you know, breathtaking, amazing, interesting, but you feel like it's a bit too much. You cannot use it, it's too graphic. This is always an ongoing debate in the news media, like which images to use, especially when we're talking about human tragedy. How far should we go?

00:27:08 NEIL CAMILLERI
Should we show the body? Not show the body, you know, so that's always, always one of the big problems. In my opinion, these news platforms that are the most likely ones to engage in the practise of emojis, I think maybe they don't put so much thought process in it. It's just a matter of, you know, choosing a base image, paste an emoji on it, let's go. Don't think about it.

00:27:43 NEIL CAMILLERI
Whereas in the mainstream media or traditional news media, there's a lot more thought given through the image selection.

00:27:51 ISAAC MUSCAT
Okay, so finally we'll discuss the future of images and inset emojis in images. Do you think earlier you mentioned how you think that this was simply just a phase where inset emojis were used. Do you think this will be the case for the time being and even the future? Or do you think there might perhaps be a time where they become popular again?

00:28:19 NEIL CAMILLERI
Look, locally, it was clearly, there was clearly a decline in the past few months with their use. They're still in use. I'm not saying they're not popular anymore, but maybe less popular than before. Whether they will come back or not, I am not sure. What I am sure of is that if we don't get emojis back, we're going to get something else.

00:28:41 NEIL CAMILLERI
Unfortunately, one of the toxic traits of the media is this copycat culture where someone starts something new and everyone starts doing the same. And this goes not only for images, but also videos, also editing styles, writing styles, and also more importantly, the subjects that we talk about. Local news is so saturated that basically if you go on any given day on all the news platforms, you're gonna find the same story in one way or another written in all of them. Even when it comes to homegrown stories. For example, a recent example, the barriers around parliament.

00:29:22 NEIL CAMILLERI
One newsroom started talking about it, then another one, and it made it their own story and their own campaign. And then someone else jumped on it. And then when finally changed, everyone was claiming victory. Oh look, it was thanks to us, it's kind of this thing.

00:29:37 NEIL CAMILLERI
With emojis I think it started as the same thing. One newsroom started doing it, then another, then another, and then pretty much all the markets started doing it. I honestly believe that we will see something maybe a bit different. I'm not sure what, but I'm pretty sure that it's gonna be something different and everyone is gonna, or almost everyone is gonna start trying to emulate it as well. Because it's, like I said, cultural competition, if you think it's something, is working for your competition, then might as well try it for yourself and keep up in the game.

00:30:11 ISAAC MUSCAT
So you mentioned how there's this copycat, like almost tradition, would you say, in the media industry. Are there any recent trends, perhaps other than set emojis that are currently emerging and you think will become more popular?

00:30:33 NEIL CAMILLERI
TikTok, for example. TikTok was something that, until a few months years ago, was something newsrooms were not really looking at. Because TikTok, in its nature, it's

childish, it's short, there's no room for any content of value in it. But slowly, slowly, I think some newsrooms started to pick it up one after the other, obviously in keeping with the TikTok style. They're trying to do these short, viral, funny videos.

00:31:09 NEIL CAMILLERI
In my opinion, it doesn't work. In my opinion, it's worse than emojis, to be honest, because it kind of belittles the work of journalists. I understand that the aim is not to be cool or something. The aim is to reach a bigger audience. And you might believe that a lot of young people use TikTok, so maybe you can reach them over there.

00:31:29 NEIL CAMILLERI
I don't think people are on TikTok for news, to be honest, because this is another thing. You need to see why people are using certain platforms. If people are not using platform to seek news, if they're only using it to seek entertainment, for example, there's no sense in trying to spoon feed or shove news down the throat on this particular platform. I think some platforms work for news, others simply don't. And I think TikTok is one of them.

00:32:02 NEIL CAMILLERI
Now, apart from emojis and TikTok, I think the other one is. What I mentioned before about the subject is that everyone's basically jumping on the same story and everyone claiming originality about the story. And then if something changes, everyone claiming victory about this change.

00:32:23 NEIL CAMILLERI
It's a bit frustrating because it shows a lack of creativity, a lack of originality, in my opinion. I experience this a lot of times as a journalist when you're working on a story, you publish some breaking story, and then next day, everyone is trying to outdo you on your own story. And it's like this unofficial competition on who's gonna take over and almost put you to shame. Like, look, I took your story and I got something much, much better than you. Now it's mine. Forget it. I have more sources, blah, blah, blah.

00:32:54 NEIL CAMILLERI
So there is this copycat culture in parts, comes from this kind of competitive attitude that exists. Like journalists, we have big egos, unfortunately, and we like to jump on the bandwagon. And if we see that someone has a story that's working, that's doing really well, maybe I should do it as well.

00:33:25 NEIL CAMILLERI

Maybe I can do it better than them. Whereas what we should be doing is either collaborating with that newsroom or trying to find something original, something different, instead of just doing what others are doing. And the problem is that if you follow someone else's story because the story is interesting, it's maybe acceptable. But if someone starts a trend like this emoji trend, which, in my opinion, I don't like at all, and everyone starts copying it. So this copycat culture can go both ways, really.

00:34:02 NEIL CAMILLERI
It can go in the right direction sometimes, because a subject might get even more coverage, maybe nationwide coverage, but then trends like emojis, if everyone starts doing it, I think the whole media sector is gonna start making a fool out of itself, to be honest, which is something that we've already done a lot with through other to other things, and we cannot afford to do it anymore.

00:34:30 ISAAC MUSCAT
So that just about concludes the interview. Thank you, Neil.

00:34:34 NEIL CAMILLERI
Welcome.

Name: Neil Camilleri

Signed: NCamilleri

Appendix C

C.1 Code Repository Link

A link to the code's respository can be found **here**.