Analysing and Detecting Extreme-Selfie Images Using Ensemble Technique

Al Amin
Energy and Electronic Engineering
University of portsmouth
Portsmouth, United Kingdom
0000-0002-2059-2352

Dr. Hongjie Ma
Senior Research Fellow
University of Portsmouth
Portsmouth, United Kingdom
0000-0001-8507-3636

Romana Alam
Faculty of Technology
University of portsmouth
Portsmouth, United Kingdom
0000-0001-5025-1025

Nasim Ahmed Roni
Research and Development
Time Research and Innovation
Dhaka, Bangladesh
nasimahmedroni@gmail.com

Md. Shazzad Hossain
Research and Development
Time Research and Innovation
Dhaka, Bangladesh
0000-0001-9507-5395

Erfanul Haque
Research and Development
Time Research and Innovation
Dhaka, Bangladesh
ehaquedipto@gmail.com

Alif B Ekram
Research and Development
Time Research and Innovation
Dhaka, Bangladesh
alifbekram@gmail.com

Redwan Abedin
Research and Development
Time Research and Innovation
Dhaka, Bangladesh
redwanabedin@gmail.com

Shah Siddiqui
Faculty of Technology
University of portsmouth
Portsmouth, United Kingdom
0000-0003-3958-9572

***Abstract*—Most individuals, especially young people, are obsessed with taking and sharing selfies online. In the age of TikTok, Facebook, and Instagram, people earn money with exceptional images or videos. However, the competition to attract viewers is not kept safe by the profilers. To some extent, people put themselves in harm's way to pursue a perfect selfie shot, causing cases of getting hurt or even dying while taking selfies. They take selfies in dangerous locations of mountain peaks, tall buildings, dangerous wild animals, lakes and many other places, which leads to many accidents. Therefore, it is a tentative proposition for the research community to understand the diverse effect of social media. This paper distinguishes between Selfies and Extreme-selfie images to detect risky situations by analyzing the surrounding. We have observed various previous Artificial Intelligence classical techniques in improving automated and accurate solutions for image classification. Additionally, we have used ensemble techniques including VGG16, VGG19, InceptionV3, ResNet50, MobileNetV2, and DenseNet121 models for extreme-selfie identification. It gives predictions based on other algorithms' results through an average voting classifier method and has shown significant success in classifying extreme-selfie images. Therefore, it outbid all other previous work achieving a validation accuracy of 97.96% and a test accuracy of 98%.**

***Index Terms*—Artificial Intelligence; Image Classification; Deep Learning; Ensemble Technique; Selfies; Extreme-Selfie Images.**

I. Introduction

A selfie is a front-facing camera photo of oneself [1]. The Oxford Dictionary's 2013 word of the year, the selfie, has brought popularity to social media with front-facing cameras [2]. Google predicted 24 billion selfies would be posted to Google Photos in 2015 [3]. According to recent research, young adults post perilous selfies online featuring wild animals, tall structures, mountains, and the ocean [4]. From 2014 through December 2016, 137 people were reportedly killed while attempting to capture a selfie [3]. Therefore, this extreme selfie capturing of such individuals leaves them through severe accidental occurrences. According to BBC News, three selfie-related fatalities were reported in 2011, 98 and 93 in 2016, and 93 in 2017. Selfies aren't typically included as a cause of death, so that the actual figure may be far higher. The absolute seriousness of the issue is understated. Selfie-related fatalities are rising in Italy, particularly among men aged 23 to 25 [5].

India, the US, and Russia are the countries with fatal selfie accidents [6]. Major killers include drowning and transportation [7]. A small but significant portion of burn trauma cases is caused by electrical injuries [6], [8]. Younger people experience more accidents and casualties. Many teens are negatively impacted by the selfie culture [7]. Therefore, to create effective interventions, it is crucial to evaluate the burden, causes, and reasons behind deaths related to taking selfies. Researchers from the US National Library of Medicine advise creating "no selfie zones" to lower fatalities. In addition, Russian authorities created public posters alerting people to the dangers of taking selfies, and Indian authorities, particularly the Mumbai Police and the Indian Railways [9], have issued alerts about potentially hazardous locations [3], [10].

Traveling to fit in on social media is common among travelers [11]. According to the authors, tourists who pass away while taking selfies differ from locals [7]. Travel and cell phone companies should encourage legal and moral self-photography [12], [13]. The author wants to estimate how many medical and nursing students in India take risky selfies. One in ten nursing and medical students had taken an unsafe selfie, and one hundred had been hurt [14]. According to additional research, social media platforms should be restricted, and mobile apps should be developed to increase dataset size by identifying extreme selfie examples online.

Globally, the practice of extreme selfie-taking and social media posting is expanding, but there is currently no cure. Our main objective is to use hybrid machine learning and deep learning techniques to examine the function and effectiveness of image-based classification. This study may serve as a warning to those who take selfies about the risks involved—aid social media companies in removing offensive or misleading content. The author measured and identified harmful selfies using Twitter articles and selfies. The model uses text, images, and geo-location information to identify potentially dangerous selfies. The selfie was examined using deep learning. However, the model's precision was poor [3], [15].

The main contributions of this paper can be summarized as follows:

1) This study used the model checkpoint saving and ensemble techniques to reduce the risk of overfitting.
2) We train our machine learning model over many iterations to avoid overfitting (50 in this case).
3) Our model assumes that overfitting will occur and will learn noises to utilize as a buffer around the equation. In this case, to avoid overfitting, the algorithm evaluates the validation accuracy and stores it if it is higher than in earlier epochs.
4) In the end, ensemble techniques improve performance and lower the risk of overfitting. We have previously tended to favor specific techniques.

The remainder of the paper is put together in this manner. The background study of this research is covered in depth in Section two. In Section three, the research methodology is provided. It consists of the dataset, data pre-processing, methodologies, and technological contribution. In addition to the result analysis, Section four includes the classification report, model evaluation, and model comparison. In Section five, the conclusion and future recommendations are given.

II. Related Works

Using the object recognition technique within an image, it is possible to determine whether or not a photograph is a selfie. An image's foreground and background features should be separated for this purpose. The author Madhuri et al. [16] investigated that the object (i.e., the person taking the selfie) can be identified based on the specifics in the foreground, and the location can be determined from the background. They also stated that it is possible to detect objects, foreground, and background information using various object identification techniques, such as exhaustive search, segmentation [17], selective search, and the Gaussian mixture model. In addition, the value of the foreground and background will be compared to a predetermined threshold value, and based on the result, it will be determined whether or not an image is a selfie. In this research, we offer a technique that can be used to determine whether an image is a selfie.

Multiple facets, such as personality, psychology, and behavior, are analyzed by researchers while examining selfies [18]. The authenticity of such selfie photographs captured during the study is questionable. Thus, many researchers tried automatically identifying selfies and removing them from the database. This research shows that these kinds of studies can help shed light on several different areas of study. It also raises questions for looking at self-portraits and for future research. In terms of soft biometrics, the author Ajita et al. [19] strategies for extracting soft-biometric features from cell phone selfies. Methods can extract demographic, physical, material, and behavioral information from selfies. They also studied why soft biometrics are preferred to main biometrics in mobile devices. This would help track authenticated users and boost authentication by utilizing the chapter's user attributes.

Additionally, the researchers investigated machine learning methods for classifying selfies as dangerous or non-threatening. Jigar et al. [20] show that basic picture features are more accurate for identifying risky selfies than multi-modal ones. High precision and Recall led to 98% accuracy, up from 73%. Precision and usefulness have improved. Using simple picture attributes to identify an image reduces the need for image captions, post locations, and text. This study presents a paradigm for classifying dangerous selfies, but it must be implemented in real-world settings. More technologies are needed to discover at-risk people and provide intervention options to avoid fatalities and injuries.

Deep Convolutional Neural Networks (CNN) [21] can play a significant role in image classification, identification, localization, pattern reorganization, and segmentation. Image classification can be defined as classifying and labeling groups of pixels or vectors within an image based on predetermined rules [22]. For instance, Garuda App is built on a deep neural network [23]. While taking selfies, the software detects the amount of risk in the backdrop and informs the user as necessary. It has a run-time average of 298ms per frame and 84% accuracy. Safe Selfie App examines various criteria, such as user velocity, local emergency contact information, and geological risks, to alert the user of potential safety issues [24]. Such applications aim to limit the number of accidental injuries and fatalities caused by photographing in hazardous situations but limit the rich visual information available from the camera. Lamba et al. [14] considered deadly selfies and selfie-related accidents by examining the victims and their deaths. The authors presented a multi-modal classifier based on non-deep learning methods. These were used to determine whether or not a particular user is in a potentially dangerous situation by considering the social media post's language, image, and location. The authors whether that caption-based image features could benefit the classification process further or not.

Avinash et al. [25] use deep ensemble learning to provide a passive method for detecting face-presentation attacks. An unauthenticated user tries to defeat the system during identity verification by downloading a social network photo and conducting an imposter attack using a printout of the user's face or a mobile device image or by developing a more complex, like a video replay attack. The authors gathered a large dataset of all types of attacks to create a deep-learning model. The researchers used facial signals and environmental cues to assess whether a person is natural or an assailant. Another ensemble model for face image identification and a technique for determining the severity of acne was proposed by the authors, Hang et al. [26]. The model has two submodules: the classification module, which assesses the severity and amount of acne, and the localization module, which, with the aid of the classification module, computes detection boxes. The results and predictions match up well.

Given that acne frequently exhibits analogous geometrical patterns across several body locations (including the face and back), the proposed model may also be used to identify acne on the back, chest, and other body regions (i.e., a cone with the apex roughly in the center). By employing a Hayashi selfie, this technology may enable patients to self-test and doctors to identify acne-related issues. The ramifications of this ensemble model for medicine. In addition, Hesham et al. [27] proposed a deep learning-based method for spotting driver distractions using an ensemble of convolutional neural networks. The first publicly available dataset for driver distraction detection with more distracting postures than competing data sets is proposed in this paper. Their 90% accurate deep learning-based solution demonstrates that a classifier ensemble weighted by a genetic algorithm increases classification confidence. Convolutional neural networks that are genetically weighted are used in the technique. This study localizes the skin, hands, and faces to investigate distraction detection. This work offers a condensed ensemble that works in real-time and has an accuracy of 84.64% regarding classification.

III. Research Methodology

The project methodologies of this study are based on the following seven stages: Collecting dataset, image preprocessing, Model execution, Model Validation, Measuring performance, and Classification. The data were collected in the first stage of this research, and then the images were pre-processed. Then the ensemble techniques were used in the model selection. After that, the proposed system model execution was performed. The following steps used model validation to validate the system, and then the performance was measured. Lastly, the classification is conducted to understand the model performance.

A. Research Dataset

The Kaggle web repository provided the majority of the images for both the selfie and Extreme-selfie data sets. Images of extreme selfies w ere t aken f rom F acebook a nd Instagram. 1318 images total, 659 of which are extreme selfies and 659 of which are normal selfies, make up the collection. The entire set of data utilized in this analysis was divided into three sections, each of which represented a different proportion of the whole: 70% was used to train the model, 15% to validate it, and 15% to test it.

B. Data Pre-processing Approaches

In a collection of 1318 images, 659 are selfies, and the remaining 659 are extreme selfies. Additionally, images are multidimensional and have various extensions. We converted all images to JPEG format to train a machine learning model with the unique Image [H*W] dimension of [224, 224] pixels. All these images were shuffled to ensure that the training data was not biased. Through split folders, libraries randomly divide 70% of data for training into 922 images, approximately 15% of data for validating into 196 images, and about 15% for testing into 200 images. Ultimately, all images were normalized using the mean and standard deviation of the pretrained dataset.

C. Algorithms Selection

This section describes the concentrated models used for this investigation. Notably, we utilized various pre-trained models, including VGG16, VGG19, InceptionV3, DenseNet121, ResNet50, and MobileNetV2, as well as the ensemble method. Ensembles are a well-established method for obtaining highly precise classifiers by combining less precise ones. An ensemble methodology aims to combine multiple models to get a better composite global model with more accurate, reliable estimates or decisions and to improve prediction performance compared to a single model.

The ensemble technique that is used in this study is the boosting method. This is one of the first and simplest effective ensemble learning methods introduced via bootstrap aggregation [28]. The meta-algorithm, a particular case of model averaging, was initially designed for classification and is typically applied to decision tree models; however, it can be applied to any classification or regression model. Using bootstrap, the method generates multiple variants of a training set. Various models are trained using each of these data sets. The outputs of the models are averaged (in the case of regression) or voted upon (in the case of classification) to produce a single result. Each model's output may be regarded as a vote for classification problems, and the ensemble model returns the class that obtains the most significant number of votes. Another name for this task is hard-voting.

On the other hand, if the result of the probabilities from every class supplied by each model is averaged and the class with the most significant average probabilities is retained; this is soft voting. In the case of this study, we have chosen this average boosting method.

We have constructed four models: Ensemble-3M, Ensemble-4M, Ensemble-6M, and Ensemble-6M. For the case of Ensemble-6M, we have used all of our trained models. Their output was averaged and put into a single output line. The average boosting method for ensemble techniques used for constructing the Ensemble-6M model is shown in Fig. 1. Ensembles of subsets of the different models were also tested.

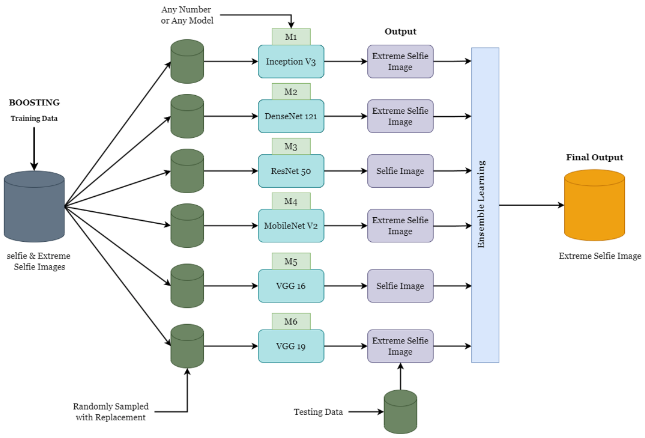

Fig. 1: Confusion Matrix of Ensamble-6M on Test Dataset

The Ensemble-3M model was built using DenseNet121, InceptionV3, and MobileNetV2 models. These three models were chosen as they performed better than all other models in this dataset. A similar method was applied for selecting the models to construct Ensemble-4M and Ensemble 5M. Ensemble-4M contains all the models from Ensamble-3M as well as VGG16. Ensemble-5M includes all the models from Ensemble-4M and VGG19.

D. Technical Contributions

This research reduced overfitting by applying the Model Checkpoint Saving and ensemble technique. Firstly, Model Checkpoint Saving can even be said to be the 'saving the best model' method here. This overfitting avoidance technique only works if the machine learning model is trained with a high epoch value, which is, in our case, 50. If our model's learning process is iterative, then there is a specific point or iteration until which the model learns the new features that we need; however, after an inevitable fact, our model will learn noises, which will result in overfitting. In that case, overfitting is avoided when the algorithm checks for validation accuracy and saves it if it is better than the previous epochs. Finally, ensemble methods increase the performance and reduce the risk of overfitting, and we have additionally favored such approaches. We have also implemented Hyperparameter tuning, which involves determining the optimal hyperparameter values for a learning algorithm and then applying this optimized algorithm to any given data set. This combination of hyperparameters maximizes the model's performance by minimizing a predefined loss function to produce more accurate results with fewer errors.

IV. Result Analysis

A. Classification Reports

This section presented the model training classification report. Numerous classification a lgorithms w ere investigated in this study. In consequence, Table 1 displays the precision, Recall, and F1 score of the classifiers, w here "P" represents precision, "R" represents Recall, and "F1" represents the F1 score. The confusion matrix assesses the precision of every algorithm. It contains False-positive, True-positive, False-negative, and True-negative values. In the Confusion Matrix, a Type 2 Error is defined as a False Negative, and a Type 1 Error is defined as a False Positive.

TABLE I: THE CLASSIFICATION REPORT ON THE DIFFERENT PRE-TRAINED MODELS (VALIDATION DATASET)

Model Name	**Extreme-selfie Class (Label-0)**			**Normal-selfie Class (Label-1)**			**Accuracy**
	P	**R**	**F1**	**P**	**R**	**F1**	
VGG16	0.97	0.96	0.96	0.96	0.97	0.96	0.96
VGG19	0.98	0.94	0.96	0.94	0.98	0.96	0.96
Inception V3	0.99	0.96	0.97	0.96	0.99	0.97	0.97
MobileNet V2	0.96	0.95	0.95	0.95	0.96	0.95	0.95
ResNet50	0.88	0.78	0.83	0.80	0.90	0.85	0.84
DenseNet 121	0.95	0.97	0.96	0.97	0.95	0.96	0.96
Ensemble3M	0.97	0.98	0.97	0.98	0.97	0.97	0.97
Ensemble4M	0.97	0.94	0.95	0.94	0.97	0.96	0.95
Ensemble5M	0.99	0.96	0.97	0.96	0.99	0.98	0.97
Ensemble6M	0.98	0.97	0.98	0.97	0.97	0.97	0.98

TABLE II: THE CLASSIFICATION REPORT ON THE DIFFERENT PRE-TRAINED MODELS (TESTING DATASET)

Model Name	**Extreme-selfie Class (Label-0)**			**Normal-selfie Class (Label-1)**			**Accuracy**
	P	**R**	**F1**	**P**	**R**	***F1***	
VGG16	0.92	0.99	0.95	0.99	0.91	0.95	0.95
VGG19	0.93	0.93	0.93	0.93	0.93	0.93	0.93
Inception V3	0.95	0.95	0.95	0.95	0.95	0.95	0.95
MobileNetV2	0.92	0.97	0.94	0.97	0.91	0.94	0.94
ResNet50	0.87	0.91	0.89	0.91	0.86	0.88	0.89
DenseNet 121	0.95	0.99	0.97	0.99	0.95	0.97	0.97
Ensemble 3M	0.95	0.99	0.97	0.99	0.95	0.97	0.97
Ensemble 4M	0.94	0.99	0.97	0.99	0.94	0.96	0.97
Ensemble 5M	0.95	0.99	0.97	0.99	0.95	0.97	0.97
Ensemble 6M	0.97	0.99	0.98	0.99	0.97	0.98	0.98

Further, Table 2 displays the findings for the testing dataset, while Table 2 in the classification report section displays the results of testing the various pre-trained models on the validation dataset. According to the categorization report for the validation dataset, there are two unique classes. Class 0(Extreme-selfie) selfies are the most severe, while Class 1(Normal-selfies) are the most prevalent. While Ensemble-3M obtains the best accuracy (98%) in the regular selfie class, the Inception V3 model has the most precision value (99%) in the extreme selfie class. On the validation dataset, we achieved a total accuracy of 98% using Ensemble-6M. For the standard selfie class, we find 99% precision using the Ensemble-4M model in table 3 of the testing dataset, while for the extreme selfie class, we get 97% precision using Ensemble-6M. On the test dataset, Ensemble-6M achieves an overall accuracy of 98%. Finally, we have enough information to prove that the Ensemble-6M performs better than all other models.

B. Model Evaluation

Fig. 2 shows the accuracy comparison of the models on the Test and Validation dataset. This illustrates the accuracy of different deep learning techniques in a distinguishable manner. In the validation dataset, the Ensamble-6M can score the highest, while Ensamble-5M, Ensemble-3M, and the pre-trained model InceptionV3 can be found very close to that. However, ResNet50 performed poorly compared with other models. This is because an ensemble model consists of two or more pre-trained models, making the output model sturdier and giving better results. In addition to this, our ensemble models are comprised of models that are already tuned for that particular dataset. This becomes a problem while performing classification tasks, with a small dataset to train on.

If the network becomes too narrow, training may be costly. In the test dataset, all ensemble models thrived, acquiring accuracy. Here, Ensemble-6M performed better than all other models when applied here. The performance of ResNet50 in the test dataset was similar to the validation dataset. While comparing both test and validation dataset accuracy scores, it can be seen that the window of highest and lowest accuracy score for the test dataset is lower than for the validation dataset.

Furthermore, most of the pre-trained models performed better on the validation dataset. But in the case of the ensemble models, it is the test dataset accuracy that came out better. Overall, Ensemble-6M from the assembled ensemble models and DenseNet121 from the pre-trained models can be considered to have performed better than others. Because the parameter configuration used in our study for that particular dataset was optimally made, the DenseNet121 pre-trained model performs better. In the case of the Ensemble-6M, this consists of all the six pre-trained models that were used on that particular dataset. So, it chose the best output out of 6 models for each classification task making this model outperform others in the process.

A visual representation of the confusion matrix in the model evaluation in Fig. 3. Ensamble-6 M model on the Test dataset forms this Confusion matrix. It shows that for regular selfies, the true positive prediction rate is 99, and the false positive prediction rate is 1, but for extreme selfies, the true positive prediction rate is 97. The false positive prediction rate is 3.

C. Model Comparison

Table 3 examines various algorithms by evaluating how effectively they foresee outcomes in light of prior research. In previous research, multiple ensemble approaches are not shown to improve accuracy to state-of-the-art levels. For example, Lamba et al. [3] get the maximum accuracy of 98% when employing a single classifier on those components. Ensemble techniques are utilized in conjunction with classifiers, such as Virk [23] and Alex et al. [29], whose results are % and 82%, respectively. These results are exciting, but there is a lack of a validation stage in the included results. These approaches aim to mix numerous models simultaneously to increase the trustworthiness of the results rather than depending solely on one model. Nevertheless, using an average vote classifier method, we created ensemble techniques that, in essence, are.

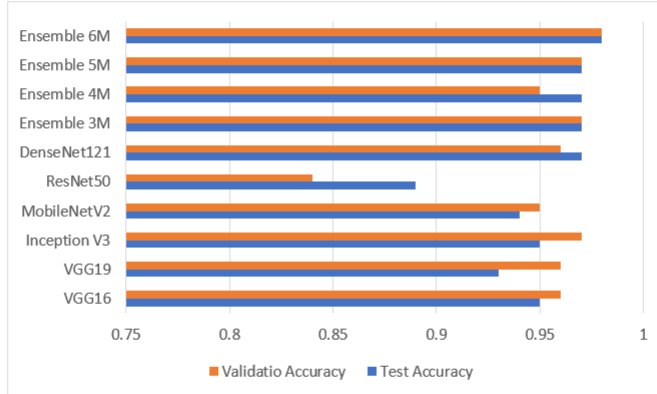

Fig. 2: Different Model Accuracy Comparison

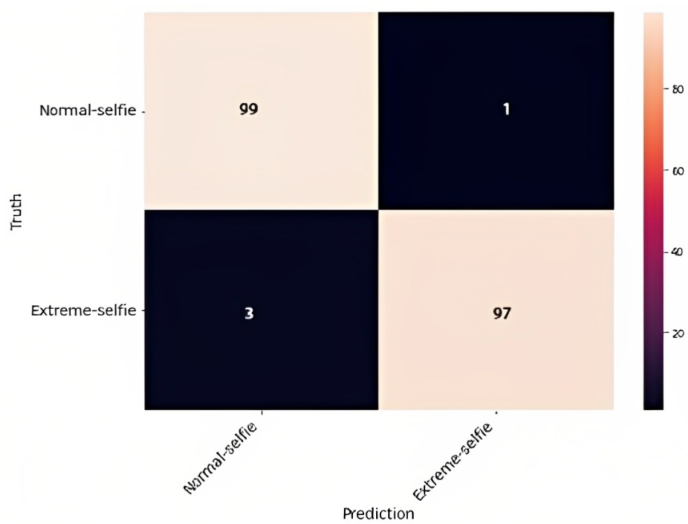

Fig. 3: Confusion Matrix of Ensamble-6M on Test Dataset

TABLE III: COMPARISONS OF PREVIOUS WORKS

Source	Model Name	Accuracy	Validation (Yes/No)
Virk et al. [23]	Deep CNN	89.28%	No
Virk et al. [23]	Separable Kernel Convolution	84.01%	No
Alex et al. [29]	SVM, MobileNetV2, VGG16	82%	No
Lamba et al. [3]	ResNet50	98%	No

statistical inferences based on the output of numerous algorithms. We tested the ensemble method and found that it worked better than other methods, with a test accuracy of 98% and a validation accuracy of 97.96%, making our model more robust than the other mentioned models.

V. Conclusion and Future Direction

While the proposed ensemble method thrived in distinguishing the Extreme selfies from the Normal ones, our work faced a significant drawback due to the dataset. Most of the data sets were collected from Kaggle, so the dataset is highly unbalanced. In the "Normal selfie" class, only a few "Group-Selfie" images can be found. This means almost all pictures contain only one "Human", not a "Group of People". In the case of the "Extreme Selfie" class, although there are not that many Group-Selfies, it is way more than in the "Normal Selfie" class. So, the model might point to the "Extreme Selfie" class, even for the "Normal Group-Selfies".

As outlined in this paper, further research has the potential to be expanded to learn more about issues associated with selfies. In addition, the precision of the results can be improved by increasing either the size of the data set or the number of annotations attached to the data set. The report's findings express optimism that future generations will use the newly acquired knowledge to develop a technological tool that warns individuals who take selfies of potentially hazardous areas and provides information about deaths that have occurred at those locations in the past. This investigation is meant to pave the way for future research and development of technological solutions that will discourage users from taking risky selfies and, as a result, reduce the likelihood of tragedies that are relatively similar to the ones that appeared here.

References

[1] K. Warfield, "Making selfies/making self: Digital subjectivities in the selfie," 2014.

[2] R. Shah and R. Tewari, "Demystifying 'selfie': a rampant social media activity," *Behaviour & Information Technology*, vol. 35, no. 10, pp. 864–871, 2016.

[3] H. Lamba, V. Bharadhwaj, M. Vachher, D. Agarwal, M. Arora, N. Sachdeva, and P. Kumaraguru, "From camera to deathbed: Understanding dangerous selfies on social media," in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 11, pp. 576–579, 2017.

[4] M. Adamkolo and H. Elmi-Nur, "Communicating 'the self'through digital images: gender bias and mental health risks associated with selfie use on social network sites," *Glob. Media J. Malays. Ed*, vol. 5, pp. 16–36, 2015.

[5] S. Gioia, D. Mirtella, L. Franceschetto, M. Lancia, F. Suadoni, and M. Cingolani, "Media-based research on selfie-related deaths in italy," *The American Journal of Forensic Medicine and Pathology*, vol. 41, no. 1, pp. 27–31, 2020.

[6] R. W. Byard, "Forensic features of fatal self photography or "selfies"," 2019.

[7] M. Linares, L. Santos, J. Santos, C. Juesas, M. Górgolas, and J.-M. Ramos-Rincón, "Selfie-related deaths using web epidemiological intelligence tool (2008–2021): a cross-sectional study," *Journal of travel medicine*, vol. 29, no. 5, 2022.

[8] D. Cuculic and I. Sosa, ""selfie"-related electrocution," *Forensic Science, Medicine and Pathology*, vol. 15, no. 4, pp. 635–637, 2019.

[9] E. Hossain, A. Alazeb, N. Almudawi, S. Almakdi, M. Alshehri, M. G. G. Faruque, and W. Rahman, "Forecasting mental stress using machine learning algorithms,"

[10] S. Bhogesha, J. R. John, and S. Tripathy, "Death in a flash: selfie and the lack of self-awareness," *Journal of travel medicine*, vol. 23, no. 4, p. taw033, 2016.

[11] J. Naslund, K. Aschbrenner, L. Marsch, and S. Bartels, "The future of mental health care: peer-to-peer support and social media. epidemiol psychiatr sci. 2016 apr; 25 (2): 113–22. doi: 10.1017/s2045796015001067."

[12] S. Jatschka, "The travel selfie: Exploring the writer's vision in contemporary travel writing," *TEXT*, vol. 23, no. Special 56, pp. 1–14, 2019.

[13] K. Leder, R. Steffen, J. P. Cramer, and C. Greenaway, "Risk assessment in travel medicine: how to obtain, interpret, and use risk data for informing pre-travel advice," *Journal of Travel Medicine*, vol. 22, no. 1, pp. 13–20, 2015.

[14] P. Behera, A. K. Singh, V. Bhatia, P. Preeti, R. Kumar, S. Das, R. Tholia, R. Ghosh, S. Kumar, K. Safiya, *et al.*, "Prevalence and determinants of the dangerous selfie among medical and nursing students: a cross-sectional study from eastern india," *BMC public health*, vol. 20, no. 1, pp. 1–7, 2020.

[15] H. Lamba, V. Bharadhwaj, M. Vachher, D. Agarwal, M. Arora, and P. Kumaraguru, "Me, myself and my killfie: Characterizing and preventing selfie deaths," *arXiv preprint arXiv:1611.01911*, 2016.

[16] M. A. Bhalekar, M. V. Bedekar, and S. Aslam, "An approach to detect an image as a selfie using object recognition methods," in *Cognitive Informatics and Soft Computing*, pp. 111–120, Springer, 2019.

[17] M. E. Hossain, M. W. Rahman, M. T. Islam, and M. S. Hossain, "Manifesting a mobile application on safety which ascertains women salus in bangladesh," *International Journal of Electrical and Computer Engineering*, vol. 9, no. 5, p. 4355, 2019.

[18] M. E. Hossain, M. Rahman, K. M. Qaiduzzaman, A. K. Shakir, and M. M. Hassan, "Efficient anti-kidnapping and anti-harassment (avoidance-detection-notification) mobile application for unwanted incidents," in *2019 IEEE Student Conference on Research and Development (SCOReD)*, pp. 112–116, IEEE, 2019.

[19] A. Rattani and M. Agrawal, "Soft-biometric attributes from selfie images," in *Selfie Biometrics*, pp. 213–225, Springer, 2019.

[20] J. Bhatt and M. Bedekar, "A survey on recognizing and significance of self-portrait images," in *International Conference on Information and Communication Technology for Intelligent Systems*, pp. 127–138, Springer, 2020.

[21] M. Elias Hossain, M. Shazzad Hossain, S. A. J. Selim Hossain, S. Alsharif, O. S. Faragallah, M. M. Eid, and A. N. Z. Rashed, "Brain tumor auto-segmentation on multimodal imaging modalities using deep neural network,"

[22] W. Rawat and Z. Wang, "Deep convolutional neural networks for image classification: A comprehensive review," *Neural computation*, vol. 29, no. 9, pp. 2352–2449, 2017.

[23] J. S. Virk and A. Dhall, "Garuda: a deep learning based solution for capturing selfies safely," in *Proceedings of the 24th International Conference on Intelligent User Interfaces: Companion*, pp. 43–44, 2019.

[24] M. Blair, D. Jeffords, E. Lilling, and S. Banik, "Safe selfie," in *2020 International Conference on Computational Science and Computational Intelligence (CSCI)*, pp. 163–169, IEEE, 2020.

[25] S. Shekhar, A. Patel, M. Haloi, and A. Salim, "An ensemble model for face liveness detection," *arXiv preprint arXiv:2201.08901*, 2022.

[26] H. Zhang and T. Ma, "Acne detection by ensemble neural networks," *Sensors*, vol. 22, no. 18, p. 6828, 2022.

[27] Y. A. Abouelnaga, M. H. S. Saad, H. M. E. Eraqi, and M. M. Moustafa, "Driver distraction identification with an ensemble of convolutional neural networks," 2019.

[28] A. J. Ferreira and M. A. Figueiredo, "Boosting algorithms: A review of methods, theory, and applications," *Ensemble machine learning*, pp. 35–85, 2012.

[29] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.