IRSTI 28.23.01

10.51889/2959-5894.2025.89.1.018

Sh. Mussiraliyeva[1], M. Bolatbek[1], Sh. Temirgaziyeva[1*]

[1]Al-Farabi Kazakh National University, Almaty, Kazakhstan
[*]*e-mail: temirgazievash@gmail.com*

DEVELOPMENT OF A METHOD FOR GRAPHICAL DESTRUCTIVE CONTENT ANALYSIS

Abstract

Social media platforms play an important role in facilitating the spread of extremism by influencing people's views, opinions, and perceptions. These platforms are increasingly used by extremist elements to spread propaganda, radicalize and attract young people. Therefore, research on the detection of extremism on social media platforms is necessary to prevent its consequences and negative consequences. It is very important to conduct a comprehensive and Comparative Study of data sets, classification methods, and screening methods using the internet extremism detection tool. The purpose of this research is to create a system for identifying ISIS and Al – Qaeda flags from images. CNN was implemented by expanding and training data sets using deep learning networks. In the course of this study, we developed a system for identifying the flags of extremist groups using images. In addition, a solution was found to increase the size of the data set. The data set contains 1,400 images, half of which are the" Al-Qaeda "flag and half are the" ISIS " flag. In addition, 2 convolutional and one fully connected layer were used to recognize the" ISIS "and" Al-Qaeda " flags. As the relevance of the work, it should be noted that the convolutional network model of CNN is trained using deep learning, that is, training based on neural networks. Novelty of the work: obtaining the highest value of the accuracy indicator using a new data set using the model classification method. The study is devoted to the study and application of deep learning methods aimed at solving the problems of identifying potentially dangerous information on the internet. With the proposed method, a 95% indicator of flag recognition from the image was achieved. The result of the proposed study can be of great help in recognizing ISIS and Al-qaeda flags and preventing terrorist activities..

Keywords: ISIS, Al-Qaeda, social networks, CNN, extremism, flag.

Ш. Мусиралиева[1], М. Болатбек[1], Ш. Темиргазиева[1]
[1]әл-Фараби атындағы Қазақ Ұлттық Университеті, Алматы қ., Қазақстан

ГРАФИКАЛЫҚ ДЕСТРУКТИВТІ МАЗМҰНДЫ ТАЛДАУ ӘДІСІН ӘЗІРЛЕУ

Аңдатпа

Әлеуметтік медиа платформалары адамдардың көзқарастарына, пікірлеріне және қабылдауларына әсер ету арқылы экстремизмнің таралуын жеңілдетуде маңызды рөл атқарады. Бұл платформаларды экстремистік элементтер үгіт-насихат тарату, радикалдандыру және жастарды тарту үшін көбірек қолданады. Сондықтан әлеуметтік медиа платформаларында экстремизмді анықтау бойынша зерттеулер оның салдары мен жағымсыз салдарының алдын алу үшін қажет. Интернеттегі экстремизмді анықтау құралын қолдана отырып, мәліметтер жиынтығын, жіктеу әдістерін және скрининг әдістерін жан-жақты және салыстырмалы түрде зерттеу өте маңызды. Бұл зерттеудің мақсаты – кескіндерден ISIS және Al-Qaeda туларын анықтау жүйесін құру. CNN терең оқыту желілерін қолдана отырып деректер жиынтығын кеңейту және оқыту арқылы жүзеге асырылды. Осы зерттеу барысында, кескіндер арқылы экстремистік топтардың жалауларын анықтайтын жүйені әзірледік. Сонымен қатар, деректер жиынтығының көлемін ұлғайтуға арналған шешім табылды. Деректер жиынтығында 1400 кескін бар, олардың жартысы "Al-Qaeda" жалауы, жартысы “ISIS” жалауы. Сонымен қатар, “ISIS” және "Al-Qaeda" жалауларын тану үшін 2 конволюциялық және бір толық қосылған қабат пайдаланылды. Жұмыстың өзектілігі ретінде CNN-нің конволюциялық желілік моделі терең оқытуды, яғни нейрондық желілерге негізделген оқытуды қолдана отырып оқытылғанын атап өткен жөн. Жұмыстың жаңалығы: модельді жіктеу әдісін қолдана отырып, жаңа деректер жиынтығын пайдалана отырып, дәлдік көрсеткішінің ең жоғары мәнін алу. Зерттеу ғаламтордағы ықтимал қауіпті ақпараттарды анықтау

мәселелерін шешуге бағытталған терең оқыту әдістерін зерттеуге және қолдануға арналған. Ұсынылған әдіс арқылы кескіннен туды танудың 95% көрсеткішіне қол жеткізілді. Ұсынылған зерттеу нәтижесі ISIS пен Al-qaeda жалауларын тануға және террористік әрекеттердің алдын алуға үлкен көмек бола алады.

Түйін сөздер: ISIS, Al-Qaeda, әлеуметтік желілер, CNN, экстремизм, ту.

Ш. Мусиралиева[1], М. Болатбек[1], Ш. Темиргазиева[1]
[1]Казахский национальный университет имени аль-Фараби, г.Алматы, Қазақстан

РАЗРАБОТКА МЕТОДА ГРАФИЧЕСКОГО ДЕСТРУКТИВНОГО КОНТЕНТ-АНАЛИЗА

Аннотация

Платформы социальных сетей играют важную роль в облегчении распространения экстремизма, влияя на взгляды, мнения и восприятие людей. Эти платформы все чаще используются экстремистскими элементами для распространения пропаганды, радикализации и привлечения молодежи. Поэтому исследования по выявлению экстремизма на платформах социальных сетей необходимы для предотвращения его последствий и негативных последствий. Всестороннее и сравнительное исследование наборов данных, методов классификации и методов скрининга с использованием онлайн-инструмента обнаружения экстремизма имеет решающее значение. Цель этого исследования – создать систему обнаружения флагов ИГИЛ и Аль-Каеда на изображениях. CNN был реализован путем расширения наборов данных и обучения с использованием сетей глубокого обучения. В ходе этого исследования мы разработали систему для обнаружения флагов экстремистских групп с помощью изображений. Кроме того, было найдено решение для увеличения размера набора данных. Набор данных содержит 1400 изображений, половина из которых-флаг "Al-Qaeda", а половина-флаг “ISIS”. Кроме того, для распознавания флагов ”ISIS" и "Al-Qaeda" использовались 2 сверточных и один полностью подключенный слой. В качестве актуальности работы следует отметить, что сверточная сетевая модель CNN была разработана с использованием глубокого обучения, то есть обучения, основанного на нейронных сетях. Новизна работы: получение максимального значения показателя точности с использованием нового набора данных с использованием метода классификации модели. Исследование посвящено изучению и применению методов глубокого обучения, направленных на решение проблем обнаружения потенциально опасной информации в интернете. С помощью предложенного метода удалось добиться 95% узнаваемости флага по изображению. Предлагаемый результат исследования может оказать большую помощь в распознавании флагов ИГИЛ и Аль-Каеды и предотвращении террористической деятельности.

Ключевые слова: ISIS, Al-Qaeda, социальные сети, CNN, экстремизм, флаг.

Main provisions

This research focuses on the identification of extremist symbols, specifically the flags of ISIS and Al-Qaeda, from images using deep learning techniques. A convolutional neural network (CNN) was employed, including two convolutional layers and one fully connected layer, to accurately classify the two types of flags. The dataset, consisting of 1,400 images, was expanded through image augmentation to improve model performance. As a result, the model achieved high accuracy in distinguishing between the two classes. The study demonstrates the effectiveness of CNN-based models in detecting potentially harmful content on social media and contributes to the prevention of extremist propaganda online.

Introduction

Social networking has now become a necessity for the world community. Every day, 2.5 quintillion bytes of data are processed around the world. A quintillion is a number two levels above a trillion, and a quadrillion is halfway between a quintillion and a trillion. This means that the world consumes so much data production in one day. Most of this data is known to be used for social media, email and Google activities. Internet users use up to 205 billion bytes of e-mail. Facebook collects 4.3 billion bytes of data per day, Instagram generates 3.6 billion bytes of data per day, and twitter collects up to 500 million bytes of data per day. And YouTube accesses 4 million bytes of data per day [1].With billions of registered users, social media platforms offer a wide range of outreach

activities. Hence, the mass media is convenient for the extremist group to promote their harmful ideology. These extremist groups share violent content and hateful messages to spread their agenda in radicalization, recruitment and propaganda [2]. Extremist organizations such as the Islamic State of Iraq and Syria (ISIS) and Al Qaeda currently use social media platforms to promote, radicalize and recruit young people. A social media tool used by ISIS in Syria to recruit new members from around the world. ISIS managed to recruit 20,000 foreign fighters through social networks, 20 percent of which came from the West.

As the amount of data on social media has grown exponentially, manual detection of radical content has become increasingly difficult. Therefore, an effective automatic method of identifying extremism is urgently needed. In addition, automatic detection of extremist profiles on social media has become a major interest and top priority of governments and counter-terrorism organizations in combating terrorist social media accounts. Therefore, the creation of information technology resources to identify cyberterrorists will help to counter extremism on the Internet [3].

Literature review

There are no definitions of the term "extremism" in the literature, as well as a number of consistent views or a universally accepted definition in the government [4]. This is not surprising, since there is no consensus among researchers and governments on the definition of the term "terrorism", which is a much less nuanced concept compared to the concept of extremism [5]. However, governments and other agencies have proposed definitions of extremism that have the following characteristics: first, an emphasis on the political, social and religious dimensions of extremism; second, to position the "extremist" as an individual or entity that actively opposes a context-dependent set of values and norms (e.g., democracy in the UK) [6]; and third, to recognize that to a certain extent extremism - and the related concept of "radicalization" - can lead to violence and, in extreme cases, acts of so-called "terrorism"[7].

One of the main challenges associated with defining extremism relates to the different contexts in which extremists are perceived. For example, the Ministry of the Interior of Great Britain [8]. The Counter-Extremism Strategy defines extremism as "overt or active opposition to our core values, including democracy, the rule of law, individual freedom, respect and tolerance for different faiths and beliefs". In contrast, Sharma [9] defined extremism as "... the extent to which violence is advocated against out-group members based on their affiliation to achieve [religious, political, or social] goals." In particular, the last definition differs in considering extremism as something related to violence (more precisely, terrorism). In contrast, the former definition recognizes extremist activities that do not necessarily lead to violence. Both definitions suggest that conceptualizations of extremist behavior may vary depending on the nature of the outgroup and the ingroup, as well as the normative values of the ingroup.

Given the above issues, it is clear that the diversity in available definitions of extremism reflects the complex, nuanced, context-dependent and multifaceted nature of the term. Another difficulty in defining these terms is that for some authors the term "extremism" is used synonymously with the terms "terrorism" and "radicalization", while for other authors these terms are considered to be related but separate [10]. For example, according to Sharma's definition, due to the link between violence and terrorism [11] all extremists are terrorists – or at least people who support terrorism. However, according to the Home Office definition, extremism is one of the many possible causes of terrorism. Regarding the relationship between extremism and radicalization, some scholars use the terms interchangeably, while others suggest that radicalization refers to the process by which an individual acquires behaviors, emotions, and beliefs that can be considered extremist.

Since the creation of profiles for online extremism is the main goal of this study, it is also worth paying attention to the religious, social and political dimensions of extremism. These dimensions are immediately noteworthy because in some cases individuals are radicalized along political lines (ie, hold extremist beliefs), while others are radicalized along religious lines, or even regarding social issues. Although each of these dimensions of extremism share common features (ie, because an

individual's religious, social, or political beliefs, emotions, and behaviors are "extreme"), the nature of each dimension is different [12]. M. Fernandez et al. in 2020 divided the online research related to extremism into three types: analysis, detection and prediction. This survey is about communication and the radicalization process. The authors also address the details of automatically identifying extremists and predicting content adoption. This review examines the lack of validated data, the lack of collaboration among researchers, the evolution of extremist language, and the lack of ethical perspectives in online research on identifying extremism [13].

M. Hashemi et al. in their 2019 research, presented the first attempt to use machine learning to recognize extremist images on web pages and flag them as symbolic propaganda. From August 2015 to September 2018, a CNN trained using 120,000 images with a generalization accuracy of 86% was used to classify 1.2 million suspicious VEO images. ISIS was 38 times more active in online propaganda than other VEOs. Most of the images are intended to condone acts of violence. In their research, they explain the massive increase in visual propaganda on the Internet and instructions for such attacks by ISIS. The study found that zero-padding is more efficient in terms of accuracy and time performance than interpolation for resizing smaller images to a fixed size for input into a CNN [14]. Y. Karimi et al. used the Python Selenium package to download images matching keywords. The photos are classified into several categories (for example, a news article focusing on a religious topic) [15]. Several pre-trained models from the PyTorch library in Python were compared to build the classifier, including networks with 18 and 34 layers (ResNets) and AlexNet [16]. The image classifier was trained on 963 Google Images and tested on 190 randomly selected and manually labeled ISIS photos. These test photos were distributed as 60 for the military, 75 for religion, and 55 for the news. Another 72 ISIS photos were used for testing (24 from each category). For testing, we used ResNet-18 with ten epochs and no rotation or horizontal and achieved 95.83% accuracy on the test sets [17].

Research methodology

Creating an ISIS and Al-Qaeda image recognition algorithm consists of a number of components, including:

1. A complex process of obtaining the necessary images;
2. Preparation of data;
3. Augmentation of received images;
4. Neural network training, as well as evaluation of its effectiveness and accuracy on test data;
5. A data set where images are stored;
6. Identification and classification of objects in images;
7. Get results.

Figure 1 below depicts the architecture of ISIS and Al-Qaeda image recognition system based on CNN model.

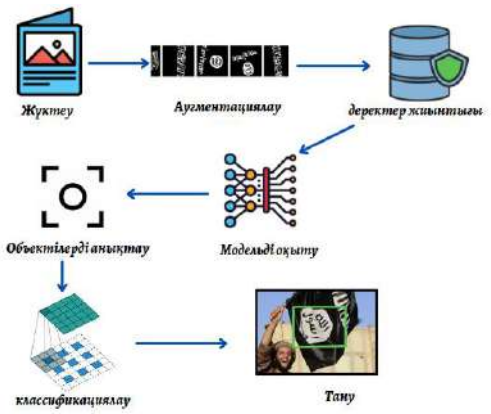

Figure 1. ISIS and Al-Qaeda image recognition system architecture

Data set. One image of the ISIS flag and one image of the Al-Qaeda flag were uploaded to develop the dataset, and data augmentation was used to generate hundreds of new ISIS and Al-Qaeda flag images based on the modifications of these two flags.

Data augmentation occurs when new data is created based on changes to existing data. In particular, for image data, data augmentation may consist of actions such as translating the image horizontally or vertically, rotating the image, zooming in or out, cropping, or changing color (Figure 2, Figure 3).

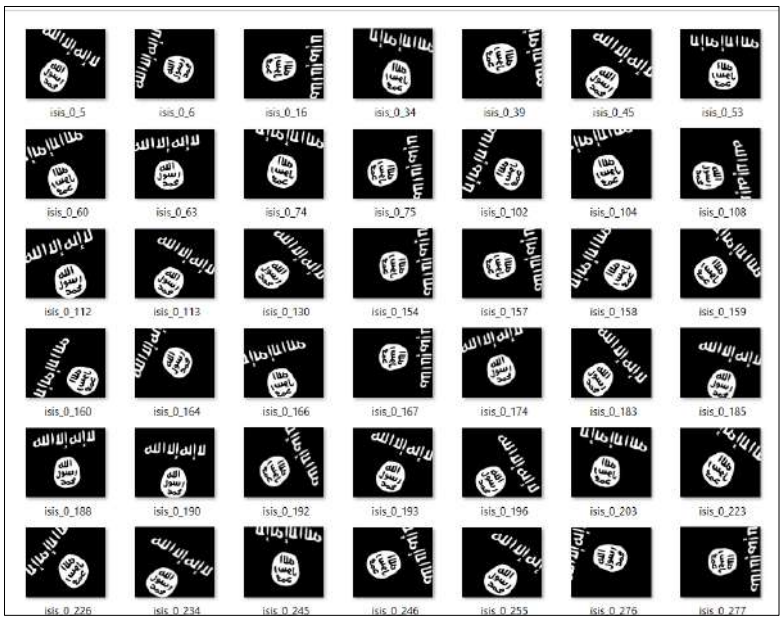

Figure 2. Magnification of ISIS image

Figure 3. Augmentation of Al-Qaeda image

Expanded image packages are compiled from the source image. The Flow() function accepts empty data and creates packets of padded data. As a result, 700 samples were obtained for each of the augmented images, totaling 1400 images (ISIS -700, Al-Qaeda -700).

The full dataset contains 1,400 images, half of which are Al-Qaeda flags and half of which are ISIS flags. The remaining script classifies 1000 samples into the training set, 200 samples into the

validation set, and 100 samples into the test set. Each set contains an equal number of Alqaeda and ISIS flags.

Neural network training. The main goal of this study is to create a system for identifying ISIS and Al-Qaeda flags from images. Data set augmentation and training were performed using image augmentation using CNN deep learning networks.

First, an image recognition model was trained using deep learning, and the trained model was tested. The deep learning process to train the model takes a lot of time, but the classification itself takes a relatively short time.

The goal of the per-sign detection problem for an image in flags is to determine whether a particular pixel belongs to a part of the sign. To solve this problem, the proposed solution is based on CNN, which is designed to classify features with geometric shapes. A sampling method is used between two fully connected layers to reduce overfitting by preventing complex co-adaptations in the training data [16]. The Keras Sequential model was used to create the CNN model. The first layer of the model is a 2D convolutional layer. In this layer, 32 output filters with kernel size 3x3 and relu activation function were used. Images are 224 pixels high by 224 pixels wide and have 3 color channels: RGB. This gives input_shape (224,224,3). Then the MaxPool2D pooling layer is added, which is designed to reduce the data size. The second Conv2D layer is performed by adding another convolutional layer with the same characteristics as the previous one, except for 64 filters. The next layer is connected with the MaxPool2D layer of the same type. The output layer of the convolution layer is Flatten and it is passed to the Dense layer. This dense layer is the eastern layer of the network, so there are 2 nodes: one for the Al-Qaeda flag and one for the ISIS flag. A softmax activation function was applied such that the output of each sample was a probability distribution over the outputs of the Alqaeda flag and the ISIS flag.

CNNs have more degrees of freedom and therefore exhibit greater variance and less bias. This fact causes CNN to overestimate the feature detection capability. Therefore, an appropriate threshold should be used. Determining precision and recall:

$$P = \frac{true\_positive}{true\_positive + false\_positive} \tag{1}$$

$$R = \frac{true\_positive}{true\_positive + false\_negative} \tag{2}$$

Then the dimension of F is:

$$F_1 = \frac{2PR}{P + R} \tag{3}$$

The threshold value used to re-estimate the final probability is determined to give the highest score in the validation data.

Results of the study

The dataset was classified using the classification function included in the scikit-learn python library. 70 percent of the data was used for training and 30 percent for testing. Accuracy functions were optimized using the Adam optimization algorithm. The proposed CNN model was tested using 10 epoch iterations.

After successfully training the neural network, it is necessary to check its performance. Figure 4 below shows a snippet of the program code. This is the stage of realization of the architecture of our model.

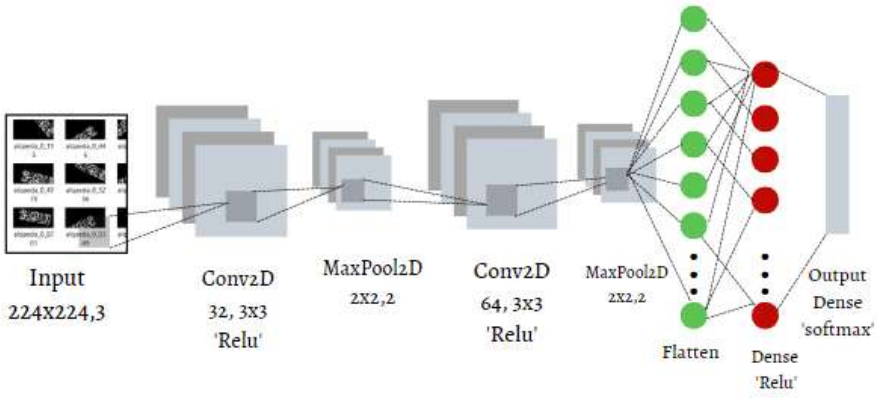

Figure 4. An example from the training phase of the CNN model

ROC curve prediction is another common tool used with binary classifiers. The dotted line shows the ROC curve of a purely random classifier, with a good classifier placing itself as far away from this line as possible (towards the upper left corner). Figure 5 below shows the accuracy of the model determined using sklearn.metrics.

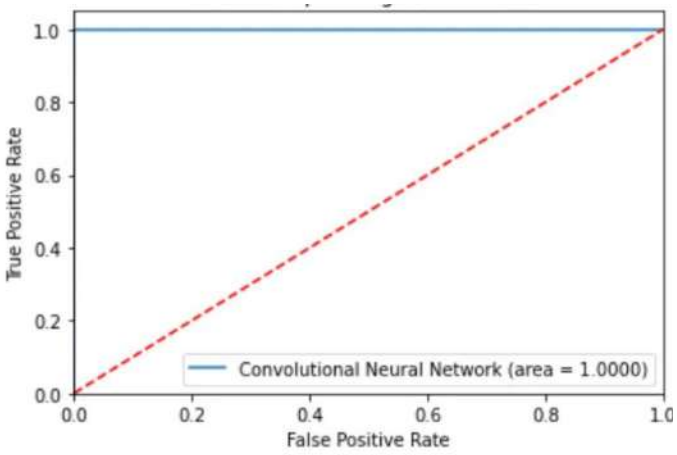

Figure 5. Model training and matching accuracy

As you can see from the images above, the neural network has been successfully trained and is ready to process new images. Figure 6 below shows an example of the feature recognition result of this system.

```
print("\033[1m The algorithm predicts, with accuracy of", accuracy, "that this is an
```

```
Found 1 validated image filenames.
The algorithm predicts, with accuracy of 100.00% that this is an image of ISIS flag
```

Figure 6. Image detected by model training

Figure 7 below shows an example of the Al-qaeda flag recognition result.

Figure 7. Image detected as a result of model training

Based on the results, 100% accuracy of ISIS and Al-Qaeda flag recognition was achieved using deep learning in the best defined category.

Discussion

The results of the conducted study showed the high efficiency of the system for recognizing the flags of extremist groups – in particular, ISIS and Al-Qaeda-based on images. The proposed method made it possible to identify destructive signs from images with an accuracy of 95% using the capabilities of deep learning. This result proves the Applied and scientific significance of identifying signs of extremist organizations through visual content.

Such a high result was achieved by increasing the data set, introducing special CNN layers (two convolutional and one fully coupled layer), as well as using focus loss.

These results also echo other studies. For example, studies based on "Multimodal Hate Speech Detection" [10] and other multimodal approaches suggest ways to detect destructive content in combination through text, image, and audio. However, our study shows that high accuracy can be achieved based only on image data. This approach is especially important in response to the tendency of extremist groups to spread their ideology visually (flag, symbol, video).

Future research should consider opportunities to increase data diversity, integrate the system with multimodal models and integrate it into real global monitoring systems. At the same time, adapting the model to work in real time will be the next step in the effective use of the system in the areas of law enforcement and cybersecurity.

Thus, the proposed solution also occupies an important place as a practical tool aimed at preventing terrorist activities on the internet.

Conclusion

The tasks set during the study were completed. A method for recognizing ISIS and Al-Qaeda flags was proposed. A solution has been found to increase the size of the dataset. To improve network performance, focus loss is used to control the network of regional proposals. Additionally, 2 convolutional and one fully connected layer were used to recognize ISIS and Al-Qaeda flags.

As the relevance of the work, it should be noted that the CNN convolutional network model was trained using deep learning, that is, based on neural networks. The novelty of the work is that the model obtains the highest value of the accuracy indicator using the new data set using the classification method.

Based on the results, 95% accuracy of ISIS and Al-Qaeda flag recognition was achieved using deep learning in the best defined category. The proposed research work can be of great help in recognizing ISIS and Al-Qaeda flags and preventing terrorist activities.

Acknowledgements

This work was funded by the AP19676342 grant of the Ministry of Science and higher education of the Republic of Kazakhstan "Multi-ideology Cyber Extremism Classification in the Kazakh language using Artificial Intelligence".

References

[1] Chaffey D. Global social media research summary August 2020, Proc. Smart Insights, 2020. ISSN: 2582-3930.

[2] Bermingham A., Conway M., L. McInerney. Combining social network analysis and sentiment analysis to explore the potential for online radicalisation, Proc. Int. Conf. Adv. Social Netw. Anal. Mining (ASONAM), 2020, pp. 231-236. http://dx.doi.org/10.1109/ASONAM.2020.31.

[3] Berger J. M. The Alt-Right Twitter Census: Defining and Describing the Audience for Alt-Right Content on Twitter, 2018. ISBN: 978-1-873769-89-8.

[4] Trip S., Bora C. H., Marian M., Halmajan A. and Drugas M. I. Psychological mechanisms involved in radicalization and extremism. A rational emotive behavioral conceptualization, Frontiers Psychol., vol. 10, pp. 437. https://doi.org/10.3389/fpsyg.2019.00437.

[5] Borum R. Radicalization into violent extremism I: A review of social science theories. J. Strategic Secur., 2020, vol. 4, no. 4, pp. 7-36. http://dx.doi.org/10.5038/1944-0472.4.4.1.

[6] Berger J. M. Extremism. Cambridge, MA, USA:MIT Press, 2019. https://doi.org/10.1080/10848770.2019.1705623.

[7] Kaya V., Tuncer S. and Baran A. Detection And Classification Of Different Weapon Types Using Deep Learning. Applied Sciences, 2021, 11 (16), 7535. https://doi.org/10.3390/app11167535.

[8] Wang K., Liu M. YOLOv3-MT: A YOLOv3 Using Multi-Target Tracking For Vehicle Visual Detection. Appl. Intell. 52, 2022, pp. 2070–2091. DOI:10.1007/s10489-021-02491-3.

[9] Manoharan S. Image Detection Classification And Recognition For Leak Detection In Automobiles. Journal of Innovative Image Processing, 01 (02), 2019, pp. 61–70. DOI:10.36548/jiip.2019.2.001.

[10] Agarwal S. and Sureka A. A focused crawler for mining hate and extremism promoting videos on YouTube. Proc. 25th ACM Conf. Hypertext Social Media, 2014, pp. 294-296. https://doi.org/10.1145/2631775.2631776.

[11] Kim J. H., Song J. H. and Lim D. CT Image Denoising Using Inception Model. Journal of the Korean Data And Information Science Society, 31 (3), 2020, pp. 487–501. https://doi.org/10.7465/jkdi.2020.31.3.487.

[12] Rink A. and Sharma K. The determinants of religious radicalization: Evidence from Kenya. J. Conflict Resolution, 2018, vol. 62, no. 6, pp. 1229-1261. https://doi.org/10.1177/0022002716678986.

[13] Fernandez M. and Alani H. Artificial intelligence and online extremism: Challenges and opportunities in Predictive Policing and Artificial Intelligence. New York, NY, USA:Taylor Francis, 2020. ISBN: 978-042-926-536-5.

[14] Hashemi M. & Hall M. Detecting and classifying online dark visual propaganda. Image and Vision Computing, 2019. DOI:10.1016/j.imavis.2019.06.

[15] Younes Karimi, Anna Squicciarini, Peter Kent Forster. A longitudinal dataset and analysis of Twitter ISIS users and propaganda. Social Network Analysis and Mining (2024) 14:19, 2024. https://doi.org/10.1007/s13278-023-01177-7.

[16] Krizhevsky A., Sutskever I. and Hinton G. Imagenet classification with deep convolutional neural networks. Advances in Neural Nnformation Processing Systems, 2012, pp. 1097–1105. https://doi.org/10.1145/3065386.

[17] Kuaban G.S., Gelenbe E., Czachórski T., Czekalski P. & Tangka, J.K. Modelling of the energy depletion process and battery depletion attacks for battery-powered internet of things (iot) devices. Sensors, 23(13), 6183, 2023. https://doi.org/10.3390/s23136183.