# Military Detection During Close Combat Situations and Border Monitoring Using Adaptive Customized Convolutional Neural Network

Vaibhav Ram S.V.N.S [1], Bhabya Sinha [1], Arunima Adhikary [1], Damodar Panigrahy [1] and Samiappan Dhanalakshmi [1]

[1] *Department of Electronics and Communications Engineering, College of Engineering and Technology, Faculty of Engineering and Technology, SRM Institute of Science and Technology, Kancheepuram, Tamil Nadu, India*

**Abstract**

In this era of 2022, we all know how much a country's relations with neighboring countries matter. A nation's security is always the top-most priority for any country and the armed forces are there to guard the borders. During any combat situation, it has been observed that the militants tend to dress like our soldiers in order to escape. This might lead to mistaking any militant for our soldiers. In order to minimize the chances of any casualty, it is very important to differentiate our soldiers from the militants. For this through our research, we propose an algorithm using the technologies of AI/ML and Deep Learning with advanced data extraction and preprocessing. An adaptive, customized Convolutional Neural Network (CNN) has been used to increase the range of prediction because we cannot go wrong while predicting as that can cause loss of life. The result is an algorithm which will flag the face detected as an Indian or foreign national.

**Keywords**

Computer vision, Military Detection, CNN, supersede network, integrated neural networks, model loading, Deep learning, Classification

## 1. Introduction

Independent light infantry forces advance into the opponent's rear sectors during battle, dodging hostile frontline strongholds and possibly isolating them for attack by later, stronger troops. Soldiers choose their own routes, targets, moments, and methods of assault; this demands a high level of expertise and training, and can be complemented by specialized equipment and weaponry to provide them with more local combat options and by taking the initiative to locate enemy weak areas. By the early modern era of warfare, defensive firepower made this strategy more and more expensive. Most such attempts were failures when trench warfare peaked in World War I [1]. Although popular and frequently effective, raiding by small parties of skilled warriors using cover and stealth proved unable to secure a resounding triumph. Terrorists frequently dress in clothes that resemble that of our soldiers or civilians in order to blend in and evade detection during combat. If our forces can't quickly identify the rebels in a crowd, they can run into difficulty in a tight location. We are working on a mechanism to provide them with information on our employees to make it simpler for them to locate the target fast and with little error because this poses a threat to innocent lives. Hence, an effective algorithm that uses the AI/ML/Deep Learning tech stack to identify nationals in conflict situations. By utilizing Image Augmentation for robustness, Convolutional Neural Networks for increased prediction range, Hyperparameter ADAM optimization for high accuracy, and High Epoch rate for better batch wise Learning, we are able to implement an approach in order to recognize or identify our countrymen using

WINS 2023: Workshop on Intelligent Systems, May 20 – 21, 2023, Chennai, India.
EMAIL: vs7961@srmist.edu.in (Vaibhav Ram S.V.N.S)
ORCID: 0009-0002-2525-2653 (Vaibhav Ram S.V.N.S)

facial recognition. This would entail choosing pertinent features, creating the network's architecture, and optimizing its settings through supervised training

# 2. Background

## 2.1. The changes were seen in the military because of AI

AI has evolved combat situations, even non-combat situations. The strategic domains have seen changes with extensive research in the field of AI. AI has existed for a long time, but proper research and development of AI in military usage has done wonders in the industry. Nuclear weapon systems have seen development with the use of AI. With time, AI's capabilities are expanding and it will become more integral to daily operations [1].

## 2.2. Evolution of creative intelligence for computer

Since hiring human role-players would be too expensive, Computer Generated Forces (CGFs) are frequently utilised in large-scale military staff exercises. The value of training as a means of preparing for actual action is increased by the use of CGFs with intelligent behaviour in exercises. Several methods were analyzed in order to enhance the efficiency of the CGFs even after the reduction of the overall cost. Even when applied to toy situations, models created using other ML algorithms where the model may be graphically represented, such as parser trees or decision trees, are difficult, if not impossible, to grasp [2].

## 2.3. Mechanisms used to analyze maritime solutions

In contexts where they must account for shifting adversarial tactics, previously unnoticed events and combinations of ordinary events hiding coordinated activity, exploitation algorithms to assist situation awareness are widely used. The system's objective is to constantly improve its ability to spot anomalies with little to no operator oversight. The system can adapt to changing circumstances while still performing effectively in contexts it has already encountered. It self-organizes to distinguish between normal and aberrant events. The latter is a critical quality: fresh learning should complement existing knowledge rather than replace it [3].

Making a learned form of scripted rule-based alerting is another use for this kind of learning. To teach models how to make the same detection choices, the rules' alerts will be used as supervised teaching signals. One benefit of this capability is that operators could change the model from the initial rules using their responses to generated alerts.

## 2.4. Usage of Bayesian Network for anomaly detection in vessel trucks

When a network is known, it is comparatively easier to analyse even after having a large number of variables. The Bayesian network (BN) was earlier used in many anomaly detection applications whereas there is no such proof of its use in maritime anomaly detection. BN has easy interpretation even without having any prior knowledge about the system. On the other hand, there have been systems which required in-depth knowledge to work upon. In this way, BN has been proven to be beneficial for researchers in many aspects [4,5].

## 2.5. Understanding the Deep Neural Networks

Deep neural networks and other machine learning methods are now essential tools for a variety of tasks like speech detection, image classification, and natural language processing. In many instances, these methods have obtained extremely high predictive accuracy on par with human performance. It was earlier believed that the simpler the model, the better the interpretability. However, the analysis is changed as now the models designed are designed in a way keeping several factors in mind which in turn makes it a complex one. These complex models have proved to be more efficient. This resulted in Deep neural networks. The DNNs are proven to have greater transparency than many other methods. Studies suggest that in the days to come, the world will experience even better working and interpretability of DNNs [5].

## 2.6. Unpaired image-to-image translation using cycle-consistent adversarial network

A class of vision and graphics problems known as "image-to-image translation" aims to learn the correspondence between an input image and an output image using a training collection of aligned image pairs. For cycle consistency, the concept of transitivity has always been in use. The method proposed can be used to make pictures from paintings, enhancing the quality of images [5,6]. One issue can be faced while training the dataset if enough characteristics are not taken into consideration.

## 2.7. Applications of artificial intelligence in military systems and its impact on citizens' sense of security

In addition to conducting research on applications in the field of civilian life, the article provides an overview of existing and anticipated prospects for the development of artificial intelligence algorithms, particularly in military applications. The application of AI algorithms in robotics, object detection, military logistics, and cybersecurity received the majority of the attention. It highlights the issues with the current fixes and how artificial intelligence could be able to help[7].

## 2.8. Framework for Federated Auto-Meta-Ensemble Learning in AI-Powered Military Operations

Federated learning, which involves collectively constructing a data pipeline, is a fantastic method for overcoming this problem. This technique works by establishing a single global model that is trained on decentralized data and applied to all users. Furthermore, the privacy and security of sensitive data handled by each institution are guaranteed by this federated arrangement. However, the usefulness and generalizability of the all-encompassing federated paradigm are seriously questioned by this process. The forecast typically has some significant biases because each machine learning system typically exhibits sensitivity in managing the given data and showing the intricate links that characterize them. This essay suggests a comprehensive federated learning strategy to deal with the aforementioned issue [8].

## 2.9. Using Intent Detection and Response Generation, Conversational AI across Military Scenarios

The development of conversational systems based on the Chinese corpus for application in military scenarios was the main goal of this endeavor. To complete their task in a strange location, the soldier will require information about their surroundings and instructions. Additionally, by deploying a conversational military agent, soldiers will get prompt, pertinent responses while working on repetitious chores with less effort and expense. In this research, a system architecture based on natural language understanding (NLU) and natural language creation is proposed for conversational military bots. (NLG). Intent detection and slot filling are the two activities included in the NLU phase. Predicting the user's intent and extracting associated entities are necessary for intent detection and slot filling [9].

## 3. Anticipated Classification Method

Here in we are proposing a solution for the detection and classification of the faces and data input, we are getting. We are using computer vision and customized neural network architecture for this problem. The most crucial part of the solution is the dataset the better it is the better the accuracy is.

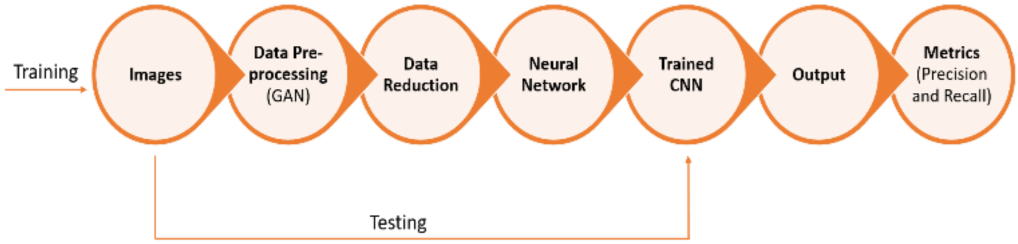


**Figure 1**: Program flow process

As shown in Figure 1, above the training dataset is loaded into the program using the classic SciPy and Spacy libraries. The dataset is then sent for cleaning in various stages, which involve the removal of background noise, corrections and additional tags. We then added filters to bring all the pictures to a recognizable and identifiable format. Once, the filters are added we normalize all the images using the augmentation techniques. Once this cleaning stage is done, we tabulate the data with the multiclass representation of every post in the dataset with 0 and 1 and so on, representing the underlying nationality of the photo shared. After tabulation, the data is passed through our neural network. The neural network is trained with the data, wherein it learns to classify the data into nationality accordingly.

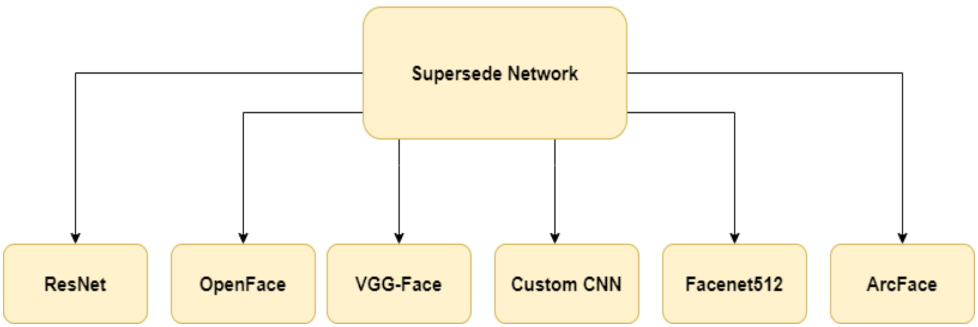


**Figure 2**: Supersede Neural Network

The neural network model we used here shown in Figure 3, is very complicated as there are multilevel integrated neural networks where our custom-made neural network is integrated into a larger neural network which has different neural networks integrated into it making it a supersede network of all the neural networks in it. Every layer in the supersede network consists of a neural network in it. The neural networks in the layers are, 'ResNet', 'YOLO', 'VGG-Face', 'OpenFace', 'Facenet', 'Facenet512', 'Deepface', 'DeepID', and 'ArcFace'. The custom-made neural network which is integrated consists of five layers, the first layer is the embedding layer which converts all the images into vectors and processes them further. The second layer consists of the activation function 'RELU' which removes the negative values in the vector to shift the whole vector grid into a positive axis, along with this a drop out of 20% is used where for every image 20% of the data will be dropped for the robustness and the rest of the image will be sent further. The next layer consists of another dropout of 20% and also a batch normalization function which brings all the values into a normalized vector above and below which is considered an outlier and won't be processed further. The fourth layer consists of another dropout of 20% and a max-pooling layer which helps in bringing and grouping all the data into processing at one time post this layer we have added in the global average pooling where it takes the values close to the average of the image vectors and then processes them further. Then comes the final layer we have the flattening layer which converts this n-dimension image into a single-dimension image that gives us the output in terms of regression result. Here, we used a random sampling technique, which involves two or more gradient descent methods along with the ADAM we chose. Once this whole process was completed, we integrated it with cloud architecture and deployed the program into any device so that our military can access it. The features are extracted using the windowing method and are precisely extracted using the harcasscade box which is one of our special features. The harcasscade is an HTML-made file which filters all the background noise and makes a box around the parts where the skin is seen in our case the face and at times the neck of the image seen.

## 4. Cloud Architecture

To enhance and deploy our solution we integrated an appropriate cloud architecture into the solution. A suitable cloud architecture can help us cut costs and increase the functionality of the solution developed. For this, we had decided to use the Docker container approach instead of using the Kubernetes approach. Docker is essentially a toolkit that enables developers to build, deploy, run, update, and stop containers using simple commands and work-saving automation through a single API [10]. Docker is a fast platform and can enable the smooth deployment of software within containers.

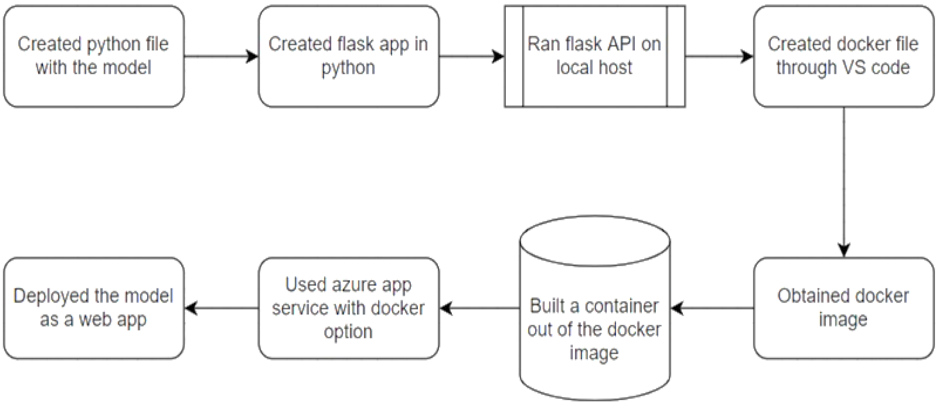


**Figure 3**: Cloud deployment process

As shown in Figure 3, we first created a Python file that consisted of the model. Using this file, we created a basic Flask Python application We ran the Flask API on the localhost to check its working constraints and functioning. A Docker file was created in VS code and was used to get the endpoints, which in our case includes the Docker images used to build a container. Once the container creation process was established, we integrated Docker with Azure in order to run Docker commands in cloud-native applications. We then obtained the URL that contained the deployed model of the solution. The cloud architecture that we selected does not have the option of clustering nodes, as doing so is computationally heavy and can at times have some major setbacks while functioning. This is just a single deployment procedure to make the understanding better we have done it. But as this model can be downloaded and the weights and biases are stored, it can be used and deployed in any integration need not be cloud integration itself. As shown above.

## 5. Model Comparison

**Table 1:** Competitive metric analysis

| Model | Accuracy (%) | Loss (%) | Precision(%) | Recall(%) |
|---|---|---|---|---|
| DNN | 86 | 0.46 | 84 | 84.02 |
| Supersede | 94 | 0.2 | 93.74 | 93 |
| YOLO-V5 | 80 | 0.54 | 77.3 | 79 |
| RNN | 82 | 0.5 | 80 | 81 |

As shown in table-1 Our supersede has 10% more accuracy than DNN, 16% more than YOLO-V5, and 14% more than RNN [11,12].

Overall, our supersede network made is better than the other networks due to its combination of neural networks that continuously process the features and produce an increase in robustness.

## 6. Results and Discussion

From table (1) we can see that supersede performed better than the rest. Hence, we decided to go forward with supersede as it was showing better results. Fig. 4 shows the results we obtained from the algorithm when processed.

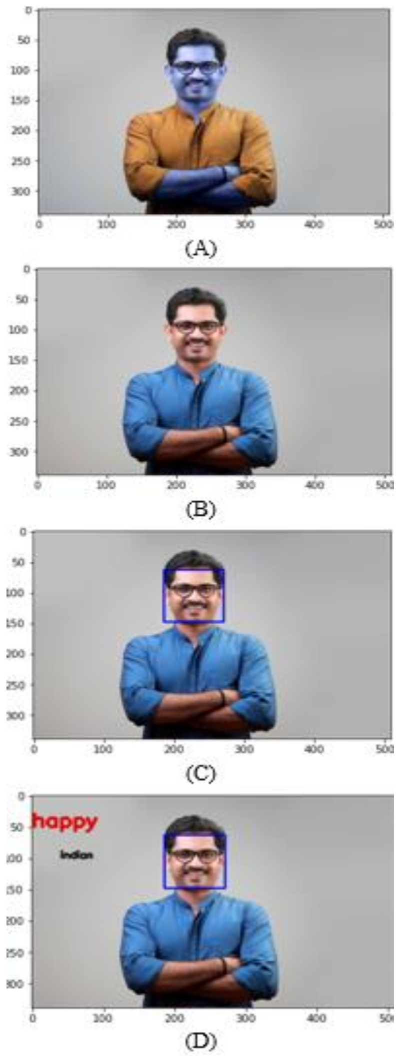


**Figure 4**: Train and Test metrics for LSTM Model: A) The input in the GBR filter, B) The input in the RGB filter, C) The input with a harcasscade, D) The input completely processed and tagged.

As we can see from Figure 4, the input image is passed through various processes to get to the final output result we want. The first major step is filtering which is done to the GBR filter see Figure 4(A), then the RGB filter is used where the whole image is brought into the natural gradient for better borders and edge of image detection without any issues [13]. Then the image gets its harcasscade which helps in removing the background noise and helps to focus on the image inside the box Figure 4(C) [14]. Once the above three steps are done then we will be able to detect the nationality and the emotion of the image for better decision making and then we send the result to the required output device Figure 4(D).

## 7. Future Enhancements

From Following are the future enhancements that could be done in order to maximize the effectiveness of this solution and help more people:

- The model can be further optimized to classify in more accurate categories.
- This model can also create its own database by storing the information processed in every single step.
- We can even make this whole process of estimation automatic where we just need to install the plugins to make it work
- Aadhar and Pan information can be linked to this algorithm database so that we can get better detection of Indians, especially in many situations along with that we can also link the criminal records in this for militant identification.

## 8. Conclusion

Thus, a technique like "Harcascades" or "Single Shot MultiBox Detector" (SSD) can be used for the face detection component to locate and identify faces in an image or video stream. To predict the nationality, race, and emotion of the subjects, a deep learning model like Convolutional Neural Networks (CNNs) could be trained on a sizable dataset of annotated face photos. This would entail choosing pertinent features, creating the network's architecture, and optimizing its settings through supervised training. It is crucial to remember that precisely identifying and categorizing demographic characteristics like race and nationality can be difficult and present ethical issues including the possibility of bias and discrimination.

The accuracy of such models might vary based on the training data and the environment in which they are utilized, and it is a challenging undertaking to identify emotions from facial expressions.

## 9. Acknowledgement

This research is done under the guidance of professors at SRM Institute of Science and Technology in the field of Applied Deep Learning. We are thankful for the guidance and the opportunity that was provided by them.

## 10. References

[1] Saalman, L.: The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk. Vol. II, Sweden: Stockholm International Peace Research Institute (2019).

[2] Wang Changqing.: The Application and Prospects of Artificial Intelligence in Cruise Missiles. http://chuansong.me/n/711504451360 (2019).

[3] Mariusz Bojarski., Davide Del Testa., Daniel Dworakowski., Bernhard Firner., Beat Flepp., Prasoon Goyal., Lawrence D. Jackel., Mathew Monfort., Urs Muller., Jiakai Zhang., Xin Zhang., Jake Zhao., and Karol Zieba.: End-to-end learning for self-driving cars. CoRR, abs/1604.07316 (2016).

[4] L. J. Luotsinen., F. Kamrani, P. Hammar., M. Ja¨ndel., and R. A. Løvlid.: Evolved creative intelligence or computer generated forces. In 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 003063–003070 (2016).

[5] Bradley J Rhodes., Neil A Bomberger., Michael Seibert., and Allen M Waxman.: Maritime situation monitoring and awareness using learning mechanisms. In Military Communications Conference, MILCOM, pp.646–652. IEEE (2005).

[6] Steven Mascaro., Ann E Nicholso., and Kevin B Korb.: Anomaly detection in vessel tracks using bayesian networks. International Journal of Approximate Reasoning, 55(1), pp.84–98 (2014).

[7] Bistron, Marta, and Zbigniew Piotrowski. "Artificial intelligence applications in military systems and their influence on sense of security of citizens." Electronics 10.7 , 871 (2021).
[8] Demertzis, Konstantinos, et al. "Federated Auto-Meta-Ensemble Learning Framework for AI-Enabled Military Operations." Electronics 12.2, 430 (2023).
[9] Chuang, Hsiu-Min, and Ding-Wei Cheng. "Conversational AI over Military Scenarios Using Intent Detection and Response Generation." Applied Sciences 12.5 ,2494 (2022).
[10] Benjamin Letham., Cynthia Rudin., Tyler H McCormick., David Madigan, et al.: Interpretable classifiers using rules and bayesian analysis: Building a better stroke prediction model. The Annals of Applied Statistics, 9(3). pp.1350–1371 (2015).
[11] Gre´goire Montavon., Wojciech Samek., and Klaus-Robert Mu¨ller.: Methods for interpreting and understanding deep neural networks. Digital Signal Processing (2017).
[12] Tim Salimans., Ian Goodfellow., Wojciech Zaremba., Vicki Cheung., Alec Radford., Xi Chen., and Xi Chen.: Improved techniques for training gans. In D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, editors, Advances in Neural Information Processing Systems 29,
[13] pp. 2234–2242. Curran Associates (2016).
[14] Jun-Yan Zhu., Taesung Park., Phillip Isola., and Alexei A Efros.: Unpaired image-to-image translation using cycle-consistent adversarial networks. arXiv preprint arXiv. Pp.1703.10593 (2017).
[15] Ashish Shrivastava., Tomas Pfister., Oncel Tuzel., Josh Susskind., Wenda Wang., and Russell Webb.: Learning from simulated and unsupervised images through adversarial training. CoRR. abs/1612.07828 (2016).