

Revealing Hidden Patterns: A Deep Learning Approach to Camouflage Detection

Rita Kamble*^{ID}, Pothuraju Rajarajeswari^{ID}

Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram 522502, India

Corresponding Author Email: reeta.kamble12@gmail.com

Copyright: ©2024 The authors. This article is published by IIETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/ijcmem.120111>

ABSTRACT

Received: 1 October 2023

Revised: 7 March 2024

Accepted: 22 March 2024

Available online: 31 March 2024

Keywords:

CNNs, LSTM, ANNs, pipeline, camouflage

In military defence and wildlife conservation operations, detecting camouflage in images poses a significant challenge. This research investigates the efficacy of deep learning techniques, including Convolutional Neural Networks (CNN), Artificial Neural Networks (ANN), and Long Short-Term Memory (LSTM), in addressing this challenge. The study meticulously evaluates each model's performance using metrics such as average accuracy, validation accuracy, and loss measures across well-known benchmark datasets comprising camouflaged and non-camouflaged images. Notably, the CNN + ANN Pipeline model emerges as the most effective, achieving a remarkable average accuracy of 91.37%. This model, together with the standalone CNN, outperforms the ANN and LSTM models in terms of camouflage detection. The discoveries advance the state-of-the-art in image analysis while also having practical implications for real-world applications. In military settings, good camouflage detection can improve situational awareness and danger detection capabilities. Similarly, automated camouflage detection helps monitor and protect endangered species by detecting hidden creatures or potential threats. Overall, this study highlights the ability of deep learning techniques to greatly improve visual analytic tasks across a variety of domains.

1. INTRODUCTION

Future generations will be based on advanced technologies such as artificial intelligence (AI), explainable AI (XAI), etc. [1, 2]. It may use unmanned systems to strike and detect the targets [3-5]. It also leads to the development of camouflage technology, in which military devices blend in with the environment so that they cannot be detected with human observation as shown in Figure 1 [6, 7]. This camouflage technique has posed significant challenges to UAV's military target striking [8, 9]. Machine learning (ML) techniques can play an important role in detecting various types of targets, and researchers have proposed numerous algorithms that use ML. However, there has been little consideration paid to military camouflaged target detection.

All of the offered ways using deep learning are grouped into two regions [10]. To begin with, prospective regions were taken into consideration such as SPP-Net, ReCNN, etc. The next one is for ideas that involve regression, like SSD, YOLO, etc. The above-mentioned standard algorithms perform badly in camouflaged target detection since they fall into the trap of the edges and backdrop of camouflaged targets because the real battlefield conditions may include desert, woods, snow, etc. It is critical in the introduction to explain why specific models such as Convolutional Neural Networks (CNN), Artificial Neural Networks (ANN), and Long Short-Term Memory (LSTM) were chosen to address the highlighted issues in camouflage target detection. These models were chosen based on their distinct capabilities and aptitude for

addressing the issues given by camouflage technology. CNNs are ideal for image processing jobs because of their capacity to capture spatial hierarchies and learn complex patterns.

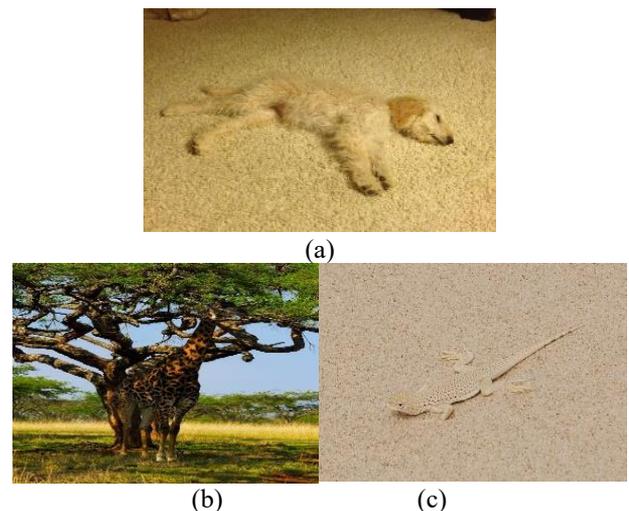


Figure 1. Camouflage images

In the context of camouflage detection, CNNs can rapidly extract features from photos, allowing the detection of small deviations and anomalies that may signal the existence of disguised targets in a variety of situations such as the desert, woodlands, or snow-covered terrain. ANNs provide flexibility in modelling nonlinear interactions within data, making them

useful for analysing complicated and heterogeneous datasets related to camouflage detection. ANN designs enable the detection system to learn complicated patterns and correlations between multiple features, improving its capacity to accurately identify camouflaged targets in a variety of backgrounds and environmental situations. LSTMs, with their ability to identify temporal dependencies in sequential data, are useful in analysing dynamic circumstances in which camouflage targets may move or alter over time. By adding LSTM networks into the detection framework, the system can successfully track and analyse temporal patterns associated with camouflaged objects, hence boosting overall detection performance in dynamic combat settings. Furthermore, it is critical to highlight how the incorporation of these unique deep learning models overcomes the shortcomings of standard algorithms, which frequently struggle with edge identification and background noise in camouflage target recognition. The suggested approach intends to solve these issues by exploiting the enhanced capabilities of CNNs, ANNs, and LSTMs, hence improving the accuracy and reliability of camouflage detection systems in real-world military applications. Incorporating these additional insights into the introduction will provide a thorough understanding of the rationale for selecting specific deep learning models and how they are expected to handle the stated issues in camouflage target detection.

1.1 Problem description

There are a number of problems that must be solved before it will be possible to military camouflaged people in images as follows:

- Due to the environment's constantly changing dynamics, it can be challenging to distinguish the texture, shape, and location of targeted camouflaged objects. For instance, the edges of targets in the desert, the desert sands, and the forest are easily absorbed into the surroundings.
- It is quite challenging to get the labelled data necessary to train and evaluate a deep learning model during the data set collection phase.
- The well-known datasets, ImageNet and VOC, cannot be used to detect and identify military targets that have been camouflaged.

1.2 Overview of proposed methodology

The Proposed method for detecting camouflage photographs makes use of a variety of deep learning models, including CNN, ANN, LSTM, and the CNN+ANN pipeline. These models were chosen for their distinct capabilities and aptitude for solving specific issues found in camouflage detecting tasks. The CNN is well-suited for extracting detailed spatial hierarchies and patterns from images, but the ANN is flexible in modelling nonlinear interactions within the data. LSTM, on the other hand, is very good at identifying temporal dependencies, which is important for analysing sequential data like image sequences or video streams. Furthermore, the CNN+ANN pipeline takes advantage of the complementing strengths of both models, combining robust feature extraction with sophisticated connection learning. The methodology seeks to solve the complexity of camouflage detection in a complete manner by utilising these several models. The models are evaluated using a variety of criteria, including average accuracy, validation accuracy, and average loss. Notably, the CNN+ANN pipeline outperforms all datasets,

highlighting its effectiveness in this domain. Furthermore, the methodology's usefulness goes beyond military and defence contexts to include a wide range of disciplines, such as wildlife conservation. This study highlights the potential of deep learning techniques to improve camouflage detection operations and establishes the framework for future advances in this sector.

2. LITERATURE REVIEW

A variety of approaches for detecting camouflage in images proposed by investigators have been discussed in this section along with the advantages over the existing terminologies [11-14].

Liu et al. [15] have proposed semi-supervised search identification network (Semi-SINet) based camouflaged military people detection system, in which, the camouflaged object detection dataset (COD10K) was taken into the consideration. It was observed that the proposed approach performed better than the existing approaches. But the obtained accuracy was low.

Ren et al. [16] suggested texture aware refinement module that separates the background in the camouflaged image, so that object can be easily detected. They applied the covariance matrices to get the texture information. In the performance analysis, COD10K and CAMO dataset were taken into the consideration. It was observed that the proposed model shows the superior performance, but still lacking to detect the dynamic camouflaged objects in the images.

Chen et al. [17] proposed the Context-aware Cross-level Fusion Network (C2F-Net), in which an information-based module was developed as a coefficient to detect camouflaged objects. A Dual-branch Global Context Module (DGCM) was also proposed to improve the informative-based features. The proposed C2F-Net performed better than state-of-the-art algorithms in the results, but a large dataset should be looked at for rigorous testing and analysis.

Lin et al. [18] addressed the issue of context aggregation strategies, where COD detection is difficult. They proposed a frequency-based context aggregation methodology known as FACA, which suppresses high frequency information. Furthermore, a gradient weighted loss function was developed to provide deep inside the contour details. During the experimentation and result analysis, it was discovered that the proposed mechanism outperforms the existing state-of-the-art techniques. But, still power of hybrid deep learning technique can be opted to enhance the accuracy.

For the COD, Fan et al. [19] proposed the SINet search identification network. They also gathered a dataset of 10,000 images known as COD10K, which serves as a benchmark dataset for testing various ML and deep learning-based models. During the experiments, it was discovered that the proposed SINet outperformed the existing works.

Shen et al. [20] proposed a COD mechanism using polarization imaging and deep learning. First, the polarization specificity is determined using the Stokes-vector-based parameter image in this methodology. The authors then suggested using the Otsu segmentation algorithm and morphological operations to highlight the target in the camouflaged-based images. In the end, 80% of the images were correctly identified, but accuracy still needs to be improved in order to be used in military and wildlife applications.

Table 1. Literature review

References	Title	Description	Advantages	Limitations
[15]	Camouflaged people detection based on a semi-supervised search identification network.	The authors presented a semi-supervised search identification network (Semi-SINet) for detecting camouflaged military personnel.	Camouflage techniques and strategies evolve with time, requiring adaptability.	Dependence on the quantity and quality of unlabeled data.
[16]	Deep texture-aware features for camouflaged object detection.	The author's suggested texture-aware refining module separates the backdrop in the camouflaged image, allowing the object to be easily spotted.	These traits are effective at distinguishing camouflaged items from complicated and variable backgrounds, hence improving detection accuracy.	The technique may be vulnerable to noise or artefacts in the input data, affecting the quality of derived features and subsequent detection accuracy.
[17]	Camouflaged object detection via context-aware cross-level fusion.	The author presented the Context-aware Cross-level Fusion Network (C2F-Net), in which an information-based module was created as a coefficient for detecting camouflaged items.	The merging of contextual information enables the model to adapt to a wide range of environmental conditions and camouflage patterns, making it suitable for a variety of scenarios.	Integrating information at several levels of abstraction complicates the detection pipeline, potentially increasing computational overhead and necessitating careful optimisation.
[18]	Frequency-aware camouflaged object detection.	The authors discuss context aggregation ways in which COD detection is challenging.	Frequency-aware approaches gather significant information across many frequency domains, resulting in a rich feature representation that improves object detection accuracy.	Extracting frequency-aware characteristics may require complex signal processing techniques, which may increase computing overhead and necessitate specialised knowledge for implementation.
[19]	Camouflaged object detection.	The author proposes. For the COD, they proposed the SINet search identification network.	The proceedings provide the most recent research and breakthroughs in camouflaged object identification, offering significant insights into cutting-edge techniques and methodologies.	The proceedings may be influenced by publication bias, in which only successful or notable research results are accepted for presentation, thus leading to an incomplete depiction of the challenges and limitations in camouflaged object detection.
[20]	Rapid detection of camouflaged artificial target based on polarization imaging and deep learning.	The authors presented a COD method that incorporates polarisation imaging and deep learning.	The combination of polarisation imaging and deep learning algorithms can considerably improve the accuracy of detecting camouflaged targets by leveraging light polarisation features that standard imaging systems may not detect.	Implementing polarisation imaging systems may necessitate specialised hardware, which can be expensive and not always available in all locations or applications.
[21]	Deep gradient learning for efficient camouflaged object detection.	The investigator proposed the COD employing the deep gradient network known as DGNet.	Deep gradient learning algorithms can successfully capture gradient-based information required for high-accuracy detection of camouflaged objects, particularly in complicated and cluttered backdrops.	Deep gradient learning approaches may have lower representation power than more complex deep learning models, resulting in inferior performance in instances involving very intricate camouflage patterns or backdrops.
[22]	Cascade and fusion: a deep learning approach for camouflaged object sensing.	The researcher suggested the COD mechanism, which solved the problem of previous algorithms, namely the difficulty in extracting informative sections such as characteristics with low signal to noise ratio.	The ability to cascade and fuse input from several stages or modalities allows the model to adapt to complicated scenarios with varying background clutter and occlusions, resulting in more accurate detection performance.	Implementing cascade and fusion techniques may increase computing complexity, particularly when combining data from various stages or modalities, thereby limiting scalability and real-time performance in resource-constrained contexts.

In the study [21], the COD was proposed by the investigator using a deep gradient network known as DGNet. Decoupling has occurred in context and a texture encoder in this mechanism. Furthermore, soft grouping chose texture and context features. In the simulation, DGNet outperforms state-of-the-art algorithms and can be used for COD after accuracy is improved.

Huang et al. [22] proposed the COD mechanism, in which

the drawback of traditional algorithms was addressed, namely, the difficulty in extracting informative parts such as features with low signal to noise ratio. They recommended the Cascade and Feedback Fusion approaches. The proposed terminology outperformed recent state-of-the-art methods in the obtained results. The state-of-Art of literature review as shown in Table 1.

2.1 Advantage of proposed scheme over existing works

List of merits of the proposed work as follows:

- The proposed deep learning models can accurately detect camouflage in images, which is an important challenge in military, defense, and wildlife conservation operations.
- The use of a combination of CNN and ANN pipeline shows promising results and can be further explored in other image detection tasks.
- The study shows that deep learning techniques can be useful in detecting camouflage in images, which can potentially save time and resources in manual detection.
- The study can contribute to the development of more effective and efficient techniques for detecting camouflage in images, leading to better security and conservation efforts.
- The proposed models are evaluated on well-known benchmark datasets, which can aid in benchmarking and comparing with other models in the future.

3. PRELIMINARIES

In this section, description of notations and abbreviations along with overview of deep learning models are presented.

3.1 Notations and abbreviations

Table 2 represents the notations and various abbreviations used in this research work.

Table 2. Description of notations and abbreviations

Abbreviations	Description
CNNs	Convolutional Neural Networks
ANNs	Artificial Neural Networks
LSTM	Long Short-Term Memory
DL	Deep Learning
ML	Machine Learning
ReLU	Rectified Linear Unit
RNNs	Recurrent Neural Networks
AI	Artificial Intelligence
Adam	Adaptive Moment Estimation
COD	Co-salient Object Detection
DGCM	Dual-branch Global Context Module
C2F-Net	Context-aware Cross-level Fusion Network
FACA	Frequency-Based Context Aggregation
DGNet	Domain Guided Network
SINet	Salient Object Detection Identification Network
VOC	Visual Object Classes
SSD	Single Shot MultiBox Detector
ReCNN	Recurrent Convolutional Neural Network
UAV	Unmanned Aerial Vehicle
XAI	Explainable AI
SPP-Net	Spatial Pyramid Pooling Network
YOLO	You Only Look Once

3.2 Overview of deep learning models

In this section brief description of DL models used is given:

3.2.1 Convolutional Neural Networks (CNNs)

In computer vision applications, notably in picture classification and object recognition, CNNs are a common

form of deep neural network. The layers of neurons that make up a CNN each learn to extract a particular set of characteristics from the input data. The input picture is typically processed by a convolutional layer, which creates a collection of feature maps by applying a number of filters. These feature maps show where specific visual patterns, such corners or edges, are prevalent across the input picture.

The pooling layer down samples the feature maps by picking the most significant values from the output of the convolutional layer. The retrieved features' dimensionality is decreased as a result of this operation, which also improves the effectiveness of the network. The final classification decision is made by a sequence of fully linked layers after the pooled characteristics have been flattened. The CNN learns its weights by a technique known as backpropagation, in which the network modifies its settings to reduce the discrepancy between the training data expected and actual labels.

3.2.2 Artificial Neural Networks (ANNs)

A group of algorithms known as ANNs are modelled after the structure and operation of the human brain. These networks may be used for a variety of tasks, including image categorization, and are built to mimic the behaviour of organic neurons. ANNs are made up of several layers of synthetic neurons that process incoming data and provide predictions. Weights that are changed during training to improve the model's performance link the layers of neurons.

ANNs are a class of algorithms that are modelled after how the human brain functions. These networks are designed to resemble the action of biological neurons and may be utilised for a range of tasks, including picture classification. Multiple layers of artificial neurons make up ANNs, which process incoming data and offer predictions. The layers of neurons are connected by weights that are altered throughout training to enhance the model's performance.

3.2.3 Long Short-Term Memory (LSTM)

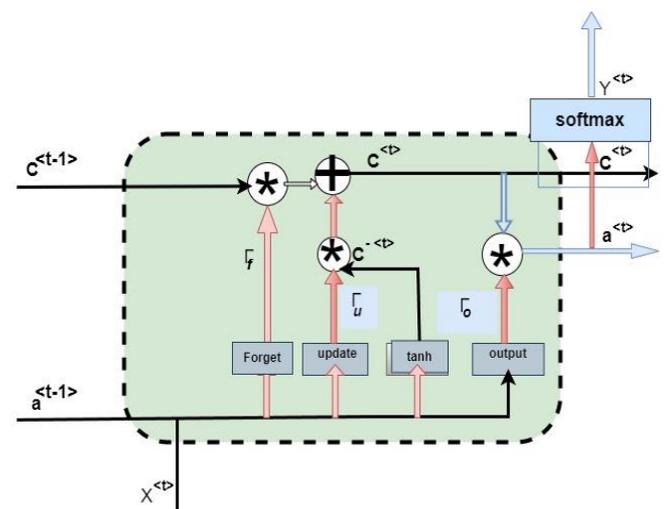


Figure 2. Pictorial representation of LSTM model

A special kind of recurrent neural network (RNN) called LSTM is made to deal with long-term dependencies in sequential input. In order to maintain information over time, LSTMs employ memory cells. This enables them to selectively forget or keep information depending on the input at each time step. The input gate, forget gate, and output gate

are three different types of gates that regulate the memory cells. The memory cell's input gate selects how much fresh input should be fed to it, the forget gate chooses which data should be removed from the memory cell, and the output gate decides how much of the memory cell should be transmitted to the following layer.

LSTMs are advantageous for tasks requiring sequential input, like as voice recognition and natural language processing, where the network must retain context from earlier portions of the sequence. This is because LSTMs employ memory cells and gates. Widely employed in both business and academics, LSTMs have been demonstrated to perform better than conventional RNNs on a number of tasks. Sequential information can be useful in finding patterns in the photos, and this research uses LSTMs for the job of detecting camouflage. The pictorial representation of LSTM circuit is demonstrated in Figure 2.

3.2.4 CNN+ANN

The combination of CNN and ANN in a pipeline is a powerful deep learning approach used for image classification tasks. CNNs are effective in extracting features from images through convolutional and pooling layers. Meanwhile, ANNs excel in making predictions based on the features extracted by the CNNs. In this pipeline approach, the output of the convolutional layers from the CNN is fed as input to the fully connected layers of the ANN, which then produce the final classification output. Thus, this combination allows the model to leverage both the feature extraction capabilities of CNNs and the prediction capabilities of ANNs, making it a powerful tool for image classification tasks.

In comparison to using CNNs or ANNs alone, the pipeline approach has demonstrated improved performance in many image classification tasks. The pipeline technique is able to discover more complex patterns in the input data, producing more accurate classification results. This is accomplished by integrating the strengths of both CNNs and ANNs. The pipeline technique is also quite adaptable and can be tailored to match various image categorization jobs.

4. PROPOSED METHODOLOGY

In this section, brief description of model architecture, Adam optimizer, categorical cross entropy loss function, training procedure and proposed algorithm is presented.

4.1 Model architecture

In this study, we trained and tested deep-learning models to classify camouflage images. The general architecture of the models consisted of a CNN followed by an ANN. The CNN portion of the model consisted of two convolutional layers, each with a 3x3 kernel size and ReLU activation function, followed by max-pooling layers with a 2x2 pool size. The output of the second max pooling layer was flattened and fed into the ANN portion of the model. The ANN consisted of three dense layers with a ReLU activation function and a final SoftMax activation function in the output layer. Dropout regularization with a rate of 0.5 was applied to the first dense layer of the ANN to prevent overfitting. The proposed model architecture is presented graphically in Figure 3.

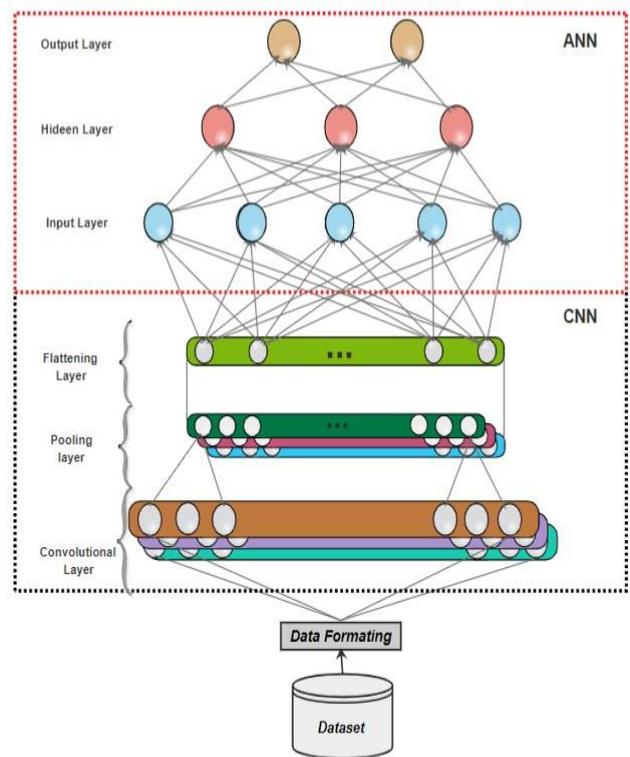


Figure 3. Proposed model architecture

The models were trained using the Adam optimizer, a stochastic gradient descent optimization algorithm that calculates adaptive learning rates for each parameter. The categorical cross-entropy loss function was used to measure the difference between the predicted and actual class labels. This is a common loss function used for multi-class classification tasks. The accuracy metric was used to evaluate the performance of the models. It represents the percentage of correctly classified images out of all images in the test set. The description of various sub-methodologies such as Adam optimizer, categorical cross-entropy and training procedure is provided in next subsections.

4.1.1 Adam optimizer

A well-liked optimisation approach called Adam (Adaptive Moment Estimation) is used in deep learning to adjust the neural network's weights as it is being trained. It is a stochastic gradient descent optimisation technique that brings together the advantages of momentum and RMSProp techniques. Based on the estimated first and second moments of the gradients, Adam adapts the learning rate for each parameter. In comparison to other optimisation techniques, this adaptive learning rate aids in faster convergence and better performance.

4.1.2 Categorical cross-entropy loss function

A typical loss function in multi-class classification issues is categorical cross-entropy. It calculates the discrepancy between the target classes' actual probability distribution and the projected probability distribution. The SoftMax activation function is used to the last layer of the neural network to produce the anticipated probability distribution. The categorical cross-entropy loss function penalises the model when it assigns a low probability to the right class in an effort to reduce the gap between these two distributions.

4.2 Training procedure

The models were trained using a stochastic gradient descent optimizer with a learning rate of 0.001 and a batch size of 32. The number of epochs varied depending on the model architecture and dataset but typically ranged from 50 to 100 epochs. Regularization techniques such as dropout and weight decay were applied to prevent overfitting.

The training and validation sets were split randomly with a ratio of 80:20, where 80% of the data was used for training and 20% was used for validation. The models were evaluated during training by monitoring the training loss and validation loss, as well as the training accuracy and validation accuracy. The categorical cross-entropy loss function was used as the primary metric for evaluating the model's performance. Additionally, the accuracy metric was used to measure the percentage of correctly classified images in both the training and validation sets. The models were trained until convergence was reached, as determined by the absence of further improvements in the validation loss and accuracy.

4.3 Performance evaluation: Accuracy

Accuracy is a metric used to evaluate the performance of a classification model. It measures the percentage of correctly predicted samples out of the total number of samples. Mathematically, accuracy is defined as the ratio of the number of correctly predicted samples to the total number of samples. It is a useful metric when the classes are balanced, meaning that each class has roughly the same number of samples.

4.4 Proposed algorithm

In this section, proposed algorithm to reveal the hidden pattern in images is presented in Algorithm 1 and its graphical flow is shown in Figure 4.

Algorithm 1: A Deep Learning Approach to Camouflage Detection

Data: Input datasets camo-covo and Camo-v

Result: Prediction accuracy of proposed deep learning model

Step 1: Split the dataset into 80:20 ratio, where 80% of dataset considered as training and rest for testing

Step 2: Apply the deep learning models, CNN, LSTM and CNN+ANN

Step 3: Monitor the models performance using training and validation loss

Step 4: Monitor the number of correctly classified images using the accuracy metrics

Step 5: Repeat step 2 to 4, until the convergence will reach

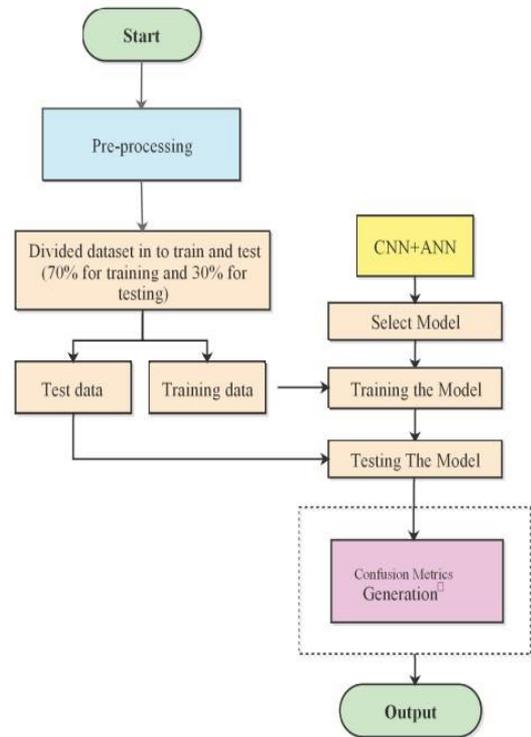


Figure 4. Graphical flow of proposed algorithm

5. RESULT AND ANALYSIS

In this section, detailed description of simulation environment and obtained results of proposed deep learning model is provided in descriptive, tabular and graphical form.

5.1 Dataset description

A camouflage dataset typically contains images of animals, insects, or other objects that have developed adaptations to merge in with their surroundings [23, 24]. It consist of 10,000 colored images. The objective of training models on such a dataset is to enable the model to detect and identify camouflaged objects in their natural environment. The dataset may contain examples of animals that utilize coloration, texture, or other physical characteristics to merge in with their environment.

A non-camouflage dataset contains images of objects that are not camouflaged and can be distinguished from their surroundings. These datasets are used to train and evaluate models for general object detection and recognition tasks, without the additional difficulty of camouflaged objects.

5.2 Experimental analysis

Table 3. The simulation parameters

Method	Dense	Max-Pooling	Convolutional	Flatten	Activation Funct.	Epochs	Batch Size	Loss Funct.	Optimizer
LSTM	2			1	Relu, Softmax	50	64	sparse_categorical_crossentropy	Adam
CNN	2	4	4	1	Relu, Softmax	50	64	categorical_crossentropy	Adam
ANN	3			1	Relu, Softmax	50	64	categorical_crossentropy	Adam
ANN+CNN	4	2	2	1	Relu, Softmax	50	64	categorical_crossentropy	Adam

The experiment involved using Python on VS Code to execute code through a Jupiter notebook extension. A computer equipped with an Intel 2.30 GHz Ryzen-7 processor and 16.0 GB of RAM was utilized for the experiment. The simulation parameters considered in the proposed work are shown in Table 3.

5.3 Results

The results obtained from the deep learning models trained on camouflage and non-camouflage images are summarized in the Table 4 and depicted in Figures 5-7. In addition, identified camouflage in images are also presented in Figure 8. The models were trained using three different architectures: CNN, ANN, and LSTM, as well as a CNN + ANN pipeline.

For the camouflage images, the CNN + ANN pipeline achieved the highest accuracy of 0.9014, while the CNN model performed the best for non-camouflage images with an accuracy of 0.9257. These results suggest that combining both CNN and ANN models can lead to better performance when detecting camouflage and non-camouflage images.

In terms of individual model performances, the CNN + ANN pipeline models generally outperformed the CNN, ANN, and LSTM models across all image types, indicating that spatial information is an important factor in detecting camouflage.

Table 4. Details of dataset with model and accuracy

Dataset	Image Type	DL Model	Average Accuracy
Camo Covo	Camouflage	CNN	0.880222146
		ANN	0.785555548
		LSTM	0.7957222031
		CNN+ANN Pipeline	0.9014444351
Non-Camouflage	Non-Camouflage	CNN	0.9146666659
		ANN	0.7979999781
		LSTM	0.7873333295
		CNN+ANN Pipeline	0.9257777863
Camo v	Camouflage	CNN	0.7853999913
		ANN	0.801000011
		LSTM	0.8063000023
		CNN+ANN Pipeline	0.9137000024

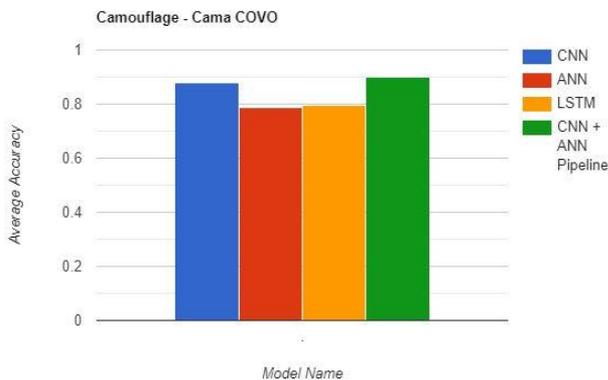


Figure 5. Obtained accuracy of DL models on camouflage in images using camo-covo dataset

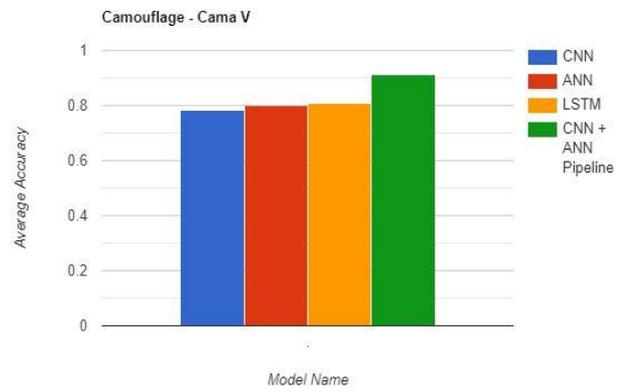


Figure 6. Obtained accuracy of DL models on camouflage in images using Camo-V dataset

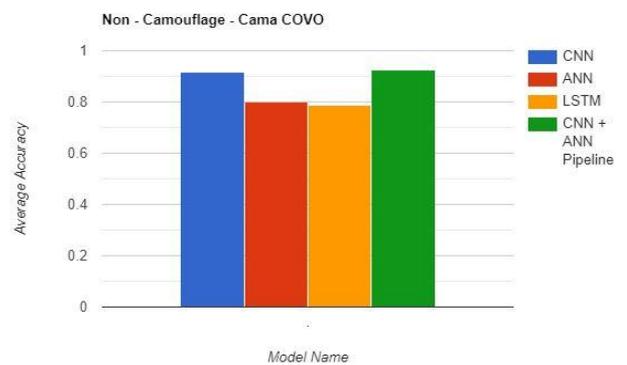


Figure 7. Obtained accuracy of DL models on non-camouflage in images using camo-covo dataset

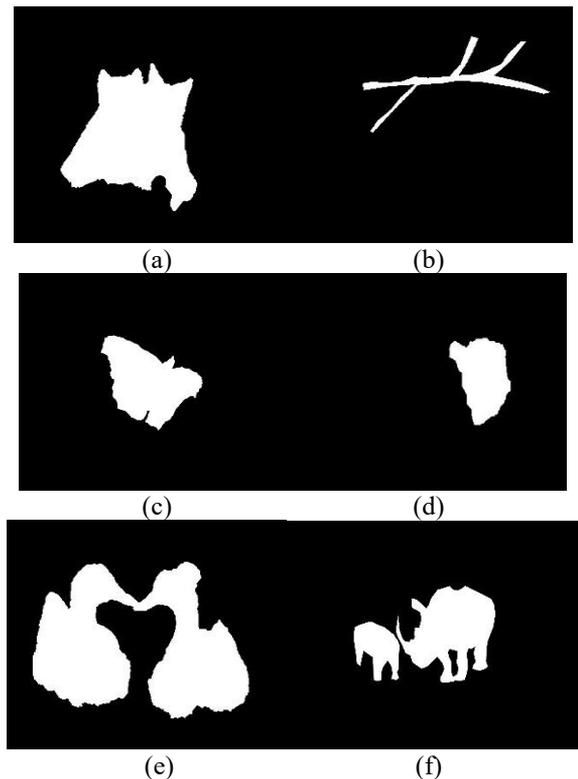


Figure 8. Identified camouflage in images

The LSTM model showed relatively lower accuracy on both camouflage and non-camouflage images, suggesting that the temporal information captured by the LSTM architecture might not be as relevant for this particular task.

5.4 Comparative analysis with existing methodologies

In this subsection, a proposed work is compared with the existing terminologies [15-17], as shown in Table 5. It was observed that the proposed deep learning model based on CNN and ANN outperformed over the recent simulated works when tested on the dataset's camo-covo and camo-v. The highest obtained accuracy to classify camouflaged in images was 91.37% in the proposed model, whereas, 90% in C² Fnet, 90.57% in TARM and 89.2% in Semi-SINet.

Table 5. Comparative analysis

Reference	Proposed Year	Methodology	Accuracy
[15]	2023	Semi-SNet	89.2
[16]	2021	TARM	90.57
[17]	2021	C ² .Fnet	90
Proposed Model	2024	CNN+ANN	91.37

6. CONCLUSION AND FUTURE RESEARCH DIRECTION

This section presents the findings and a summary of the proposed methodology, as well as the direction for future research.

6.1 Conclusion and summary

The study goes thoroughly into deep learning approaches, with a particular emphasis on Convolutional Neural Networks (CNN) and Artificial Neural Networks (ANN), which are formidable tools for handling the complex challenge of detecting camouflaged images. Through comprehensive investigation and rigorous comparison, the study reveals the extraordinary performance of the CNN + ANN pipeline model, which emerges as the unchallenged leader with the highest average accuracy across all scrutinised datasets. This robustness not only demonstrates the efficacy of the hybrid approach, which strategically combines CNN for feature extraction and ANN for classification, but also highlights its enormous potential across a wide range of practical applications, from military defence to wildlife conservation operations. While the abstract hinted at this extraordinary outcome, a more exact expression of specific accuracy values would surely strengthen the conclusions drawn and provide a more detailed understanding of the model's capabilities in compared to its competitors. Surprisingly, the accuracy of this hybrid model, CNN+ANN, is 91.37%, confirming its status as a powerful rival in the field of deep learning-based image analysis and pattern identification. This high level of accuracy not only demonstrates the model's efficacy, but also emphasises its dependability in real-world settings where precision is critical.

6.2 Future findings

Furthermore, the study suggests intriguing avenues for future research, urging the investigation of more complicated scenarios, such as real-time video surveillance, to determine

the models' flexibility and efficacy in dynamic settings. Furthermore, the optimisation of these models through the integration of varied datasets and rigorous fine-tuning of hyperparameters provides a fertile ground for improving their performance and robustness under difficult conditions. By condensing these complex insights and reaffirming major findings, the summary not only gives a thorough overview of the study's achievements, but it also lays the groundwork for future research in the field of deep learning-based image analysis and pattern recognition.

REFERENCES

- [1] Gao, S., Wu, J., Ai, J. (2021). Multi-UAV reconnaissance task allocation for heterogeneous targets using grouping ant colony optimization algorithm. *Soft Computing*, 25(10): 7155-7167. <https://doi.org/10.1007/s00500-021-05675-8>
- [2] Huang, Y., Ding, W., Li, H. (2016). Haze removal for UAV reconnaissance images using layered scattering model. *Chinese Journal of Aeronautics*, 29(2): 502-511. <https://doi.org/10.1016/j.cja.2016.01.012>
- [3] Park, J.H., Choi, S.C., Ahn, I.Y., Kim, J. (2019). Multiple UAVs-based surveillance and reconnaissance system utilizing IoT platform. In 2019 International Conference on Electronics, Information, and Communication (ICEIC), Auckland, New Zealand, pp. 1-3. <https://doi.org/10.23919/ELINFOCOM.2019.8706406>
- [4] Kim, S., Kwon, D., Ji, Y. (2019). CNN based human detection for unmanned aerial vehicle (poster). In Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services, Seoul Republic of Korea, pp. 626-627. <https://doi.org/10.1145/3307334.3328659>
- [5] Wang, Y., Luo, X., Luo, L., Zhang, H., Wei, X. (2020). UAV tracking based on saliency detection. *Soft Computing*, 24: 12149-12162. <https://doi.org/10.1007/s00500-019-04652-6>
- [6] Troscianko, T., Benton, C.P., Lovell, P.G., Tolhurst, D.J., Pizlo, Z. (2009). Camouflage and visual perception. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1516): 449-461. <https://doi.org/10.1098/rstb.2008.0218>
- [7] Sengottuvelan, P., Wahi, A., Shanmugam, A. (2008). Performance of decamouflaging through exploratory image analysis. In 2008 First International Conference on Emerging Trends in Engineering and Technology, Nagpur, India, pp. 6-10. <https://doi.org/10.1109/ICETET.2008.232>
- [8] Zhang, Y., Xue, S.Q., Jiang, X.J., Mu, J.Y., Yi, Y. (2013). The spatial color mixing model of digital camouflage pattern. *Defence Technology*, 9(3): 157-161. <https://doi.org/10.1016/j.dt.2013.09.015>
- [9] Yang, X., Xu, W. D., Jia, Q., Liu, J. (2021). MF-CFI: A fused evaluation index for camouflage patterns based on human visual perception. *Defence Technology*, 17(5): 1602-1608. <https://doi.org/10.1016/j.dt.2020.08.007>
- [10] He, K., Zhang, X., Ren, S., Sun, J. (2015). Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(9): 1904-1916. <https://doi.org/10.1109/TPAMI.2015.2389824>
- [11] Zheng, Y., Zhang, X., Wang, F., Cao, T., Sun, M., Wang,

- X. (2018). Detection of people with camouflage pattern via dense deconvolution network. *IEEE Signal Processing Letters*, 26(1): 29-33. <https://doi.org/10.1109/LSP.2018.2825959>
- [12] Fang, Z., Zhang, X., Deng, X., Cao, T., Zheng, C. (2019, May). Camouflage people detection via strong semantic dilation network. In *Proceedings of the ACM Turing Celebration Conference, China*, pp. 1-7. <https://doi.org/10.1145/3321408.3326662>
- [13] Xue, F., Yong, C., Xu, S., Dong, H., Luo, Y., Jia, W. (2016). Camouflage performance analysis and evaluation framework based on features fusion. *Multimedia Tools and Applications*, 75: 4065-4082. <https://doi.org/10.1007/s11042-015-2946-1>
- [14] Feng, X., Guoying, C., Wei, S. (2013). Camouflage texture evaluation using saliency map. In *Proceedings of the Fifth International Conference on Internet Multimedia Computing and Service, Huangshan China*, pp. 93-96. <https://doi.org/10.1145/2499788.2499877>
- [15] Liu, Y., Wang, C.Q., Zhou, Y.J. (2023). Camouflaged people detection based on a semi-supervised search identification network. *Defence Technology*, 21: 176-183. <https://doi.org/10.1016/j.dt.2021.09.004>
- [16] Ren, J., Hu, X., Zhu, L., Xu, X., Xu, Y., Wang, W., Deng, Z., Heng, P.A. (2021). Deep texture-aware features for camouflaged object detection. *IEEE Transactions on Circuits and Systems for Video Technology*, 33(3): 1157-1167. <https://doi.org/10.1109/TCSVT.2021.3126591>
- [17] Chen, G., Liu, S.J., Sun, Y.J., Ji, G.P., Wu, Y.F., Zhou, T. (2022). Camouflaged object detection via context-aware cross-level fusion. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(10): 6981-6993. <https://doi.org/10.1109/TCSVT.2022.3178173>
- [18] Lin, J., Tan, X., Xu, K., Ma, L., Lau, R.W. (2023). Frequency-aware camouflaged object detection. *ACM Transactions on Multimedia Computing, Communications and Applications*, 19(2): 1-16. <https://doi.org/10.1145/3545609>
- [19] Fan, D.P., Ji, G.P., Sun, G., Cheng, M.M., Shen, J., Shao, L. (2020). Camouflaged object detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA*, pp. 2777-2787.
- [20] Shen, Y., Lin, W., Wang, Z., Li, J., Sun, X., Wu, X., Wang, S., Huang, F. (2021). Rapid detection of camouflaged artificial target based on polarization imaging and deep learning. *IEEE Photonics Journal*, 13(4): 1-9. <https://doi.org/10.1109/JPHOT.2021.3103866>
- [21] Ji, G.P., Fan, D.P., Chou, Y.C., Dai, D., Liniger, A., Van Gool, L. (2023). Deep gradient learning for efficient camouflaged object detection. *Machine Intelligence Research*, 20(1): 92-108. <https://doi.org/10.1007/S11633-022-1365-9>
- [22] Huang, K., Li, C., Zhang, J., Wang, B. (2021). Cascade and fusion: A deep learning approach for camouflaged object sensing. *Sensors*, 21(16): 5455. <https://doi.org/10.3390/s21165455>
- [23] Le, T.N., Nguyen, T.V., Nie, Z., Tran, M.T., Sugimoto, A. (2019). Anabran network for camouflaged object segmentation. *Computer Vision and Image Understanding*, 184: 45-56. <https://doi.org/10.1016/j.cviu.2019.04.006>
- [24] Yan, J., Le, T.N., Nguyen, K.D., Tran, M.T., Do, T.T., & Nguyen, T.V. (2021). Mirrornet: Bio-inspired camouflaged object segmentation. *IEEE Access*, 9: 43290-43300. <https://doi.org/10.1109/ACCESS.2021.3064443>