Asian Journal of Technology and Applied Sciences
ISSN: 2301-xxxx (online) | IF: 10.01 | Vol: 1, issue 1
Available: http://asiansciencejournal.org

A MODEL BASED ON ARTIFICIAL INTELLIGENCE FOR EVALUATING THE EFFECTIVENESS OF MULTIMODAL CAMOUFLAGE OF MILITARY OBJECTS AGAINST UNMANNED AERIAL VEHICLE SURVEILLANCE SYSTEMS

RAXMANOV SHOHIJAHON SHARIPOVICH
Major, Senior Instructor for UAV Control and Counter-UAV Operations, Department of Advanced Military Technologies, Faculty of the Central Military District, University of Military Security and Defense of the Republic of Uzbekistan

Abstract: This article proposes an AI-based model for evaluating the detectability of objects based on visible and thermal imagery captured by unmanned aerial vehicles (UAVs). The model integrates channel-specific assessments, contextual distinctiveness, and sensor quality metrics to calculate an integral risk factor and relative change. The numerical values provided serve as an illustrative example of the model's operation and do not represent the results of practical experiments.

Keywords: unmanned aerial vehicle (UAV), artificial intelligence, multimodal imagery, visible channel, thermal imagery.

Introduction

Optical, thermal, and multispectral surveillance sensors mounted on unmanned aerial vehicles (UAVs) have expanded the capability to detect military objects across various altitudes, viewing angles, and lighting conditions. Even if an object blends into its background within the visible spectrum, its thermal signature, contours, or the logic of its placement can serve as sufficient indicators for an automated detection system. Consequently, evaluating camouflage solely based on a human observer's subjective judgment fails to fully account for the realities of modern multi-sensor surveillance.

In aerial imagery, targets often appear small, partially obscured, and situated within complex backgrounds. Factors such as camera angle, motion blur, image resolution, and the object's orientation can degrade detection performance. However, detectors based on machine learning analyze not only color but also texture, contours, thermal contrast, and semantic relationships with surrounding objects. Therefore, when assessing the effectiveness of camouflage, it is essential to consider features that may be imperceptible to the human eye but remain detectable by algorithms. It has been noted that methods for detecting anomalies in multispectral imagery can achieve an area-under-the-curve (AUC) value greater than 0.9 for certain camouflaged targets. In the case of hyperspectral sensors mounted on unmanned aerial vehicles (UAVs), selecting spectral bands tailored to the specific task helps reduce data volume and facilitates the discrimination of concealed materials. These findings indicate that camouflage must be evaluated simultaneously across color, thermal, and contextual dimensions.

The objective of this research is to develop an artificial intelligence model for UAV-based surveillance that integrates the visible, thermal, and contextual observability of a military object into a single quantitative metric. The model is intended not for the design of camouflage solutions, but for the comparison of existing solutions under standardized testing conditions. The numerical table provided in the article serves as an illustrative calculation to demonstrate the model's performance and should not be interpreted as the result of a field experiment.

Research Methodology

Asian Journal of Technology and Applied Sciences
ISSN: 2301-xxxx (online) | IF: 10.01 | Vol: 1, issue 1
Available: http://asiansciencejournal.org

The methodological framework of this research consists of methods such as systems analysis, comparative analysis, functional analysis, and generalization. These methods are employed to systematically examine the structure of the SIMNB model, the interrelationships among its indicators, and the computational results. No new field or laboratory experiments were conducted for this study; therefore, the numerical examples are interpreted as hypothetical data intended to illustrate the model's computational procedure.

1. Systems Analysis Method

Systems analysis is applied to treat the SIMNB model as an integrated assessment system composed of interconnected elements. The visible channel score (P_v), thermal channel score (P_t), contextual distinctiveness indicator (C), and other metrics representing channel quality are considered individual elements. Their combined outcome is aggregated into an integral indicator via a weighted sum: The requirement that the sum of the weights must equal one ensures a consistent interpretation of the model's output.

The application of this method involves establishing the logical connections between the input indicators, the rules for their aggregation, and the final result. Systems analysis also makes it possible to distinguish the outcome of an individual channel from an integral indicator, and to examine the contribution of each component to the result separately. Its scientific contribution lies in providing a coherent description of the model's structure, revealing the interconnections between computational stages, and establishing a theoretical foundation for subsequently validating the model using independent data.

2. Comparative Analysis Method

Comparative analysis is employed to evaluate the hypothetical scenarios A, B, C, and D—presented in the article—based on consistent assessment criteria. This comparison considers evaluations derived from both visible and thermal channels, alongside contextual indicators and the integral R_d value. A fundamental requirement here is that the computational rules and the assigned weights remain constant across all scenarios; otherwise, interpreting the differences between the results would become problematic.

Since the figures provided in the article's table represent only sample calculations, this comparative analysis should not be interpreted as empirical evidence or proof of performance under real-world conditions. Rather, they illustrate the operational mechanics of the formulas and demonstrate how various input values influence the integral indicator. The scientific contribution of this method lies in the systematic comparison of scenarios based on a unified set of criteria, the consistent interpretation of results, and the identification of key indicators that should be monitored in future experiments.

3. Functional Analysis Method

Functional analysis is used to examine the specific function assigned to each indicator within the model and its respective role in the overall assessment. Indicators from the visible channel represent visual features found in RGB imagery; The thermal channel indicator represents information based on temperature or radiation differences in an infrared image, while the contextual indicator describes the object's relationship with its environment in a general sense. Sensor quality indicators, meanwhile, are designed to account for the reliability of data from a specific channel within the current assessment.

This method provides a mathematical explanation of how the integral result changes when the value within a specific channel of the model is altered. For instance, in a weighted sum formula, the contribution of each component depends on its value and its corresponding weighting coefficient. However, this relationship alone does not guarantee the model's accuracy under real-world conditions; such assurance requires verification using independent data, calibration, and error metrics. The scientific contribution of functional analysis lies in

Asian Journal of Technology and Applied Sciences
ISSN: 2301-xxxx (online) | IF: 10.01 | Vol: 1, issue 1
Available: http://asiansciencejournal.org

isolating the role of each parameter, explaining the model's output, and clarifying the assumptions that require validation.

4. Generalization Method

The generalization method is used to synthesize conclusions derived from systemic, comparative, and functional analyses. Based on this approach, the SIMNB model is characterized as a conceptual framework designed to integrate indicators from multiple observation channels into a unified evaluation system. Furthermore, it emphasizes the necessity—for the sake of accurate interpretation of the model's output—of considering data quality, sensor calibration, the selection of weights, the partitioning of datasets into independent subsets, and the assessment of uncertainty in the results. Through generalization, the model's theoretical utility—namely, the ability to express metrics using a unified criterion while keeping individual components interpretable—is distinguished from its limitations. Specifically, the illustrative percentages presented in the article cannot be automatically applied to real-world conditions, nor are they sufficient to draw independent conclusions regarding practical effectiveness. The scientific contribution of this method lies in the cautious generalization of research findings, the grounding of conclusions within the scope of existing evidence, and the establishment of requirements for future empirical verification.

Methodological limitation. The model and numerical examples provided in the article serve to illustrate a conceptual evaluation approach. As the study lacks validation on an independent dataset, statistical confidence intervals, and replicable experiments, the results should not be regarded as a verified indicator of practical effectiveness.

Observability Factors

Three groups of features are identified for evaluation. The visible channel provides information regarding color harmony, brightness contrast, texture, and external contours. The thermal channel represents the differences in surface temperature and radiation between an object and its background. The contextual channel evaluates the object's location, size, and geometry, as well as its deviation from typical structures within the environment. Categorizing features in this way allows for the analysis of which modal channel contributes most significantly to the overall threat level.

In the context of visible imagery, the acronym "RGB" refers to an image composed of red, green, and blue color channels, while "TIR" denotes an image from the thermal infrared range. Paired RGB and TIR frames are aligned in terms of time, viewing angle, and field of view. If spatial misalignment persists, an incorrect combination of features from the two channels could compromise the accuracy of the detection assessment; therefore, the quality of the registration (alignment) is verified separately.

Evaluation Architecture

The proposed model—referred to as the "Artificial Intelligence-based Multimodal Camouflage Evaluation Model" (or "SIMNB" model for short)—operates according to the sequence illustrated in Figure 1.

Figure 1. Operational algorithm of the SIMNB model. In this context, artificial intelligence does not generate a mask; rather, it evaluates masked and unmasked states based on recurring criteria. Each channel detector returns a score between 0 and 1 indicating the presence of an object. Here, P_v represents the detection score for the visible channel, P_t for the thermal channel, and C denotes the level of contextual distinctiveness. The final risk value, R_d, is determined by their weighted sum:

$R_d = w_v P_v + w_t P_t + w_c C$ (1)

Where w_v, w_t, and w_c are the weighting coefficients for the respective channels; these are non-negative and satisfy the following normalization condition:

Asian Journal of Technology and Applied Sciences
ISSN: 2301-xxxx (online) | IF: 10.01 | Vol: 1, issue 1
Available: http://asiansciencejournal.org

$w_v + w_t + w_c = 1$ (2)

The weights are adjusted based on sensor quality. If q_i represents the current image quality score for channel i, the weights can be normalized as follows:

$w_i = q_i / (q_v + q_t + q_c)$ (3)

For instance, if fog or low-light conditions degrade the quality of the RGB channel, its weight decreases. Similarly, saturation of the thermal sensor or a background temperature approaching that of the object reduces the weight of the thermal (TIR) channel. While weights can be learned using a test dataset, it is essential to ensure that the training, validation, and final test datasets remain distinct and do not overlap.

Masking Effectiveness Criterion

Let R_{before} represent the risk before masking, and R_{after} represent the risk after masking. Provided that R_{before} is not close to zero, the relative effectiveness is calculated using the following expression:

$E_{mask} = (1 - R_{after} / R_{before}) \times 100\%$ (4)

If E_{mask} is positive, it indicates a reduction in detectability as estimated by the model; if it is close to zero, there is no significant change; if it is negative, it is interpreted as an increase in the risk of detection following masking. This value does not represent the actual probability of detection in a combat scenario. Rather, it is a relative metric dependent on the chosen sensors, the dataset, the detector, and the testing conditions.

Evaluation Procedure

The evaluation is based on pairs of images of the same object—one in an unmasked state and the other in a masked state. The image frames are stratified according to altitude, range, season, time of day, and lighting conditions. Within each stratum, the imaging protocol, sensor settings, and the criteria for manual annotation must remain consistent. The same detector and decision threshold are applied to the various masking configurations being compared.

Detector performance is assessed using precision, recall (completeness of detection), the F1-score, and mean average precision. Precision represents the ratio of true positive results to all positive detections, while recall (completeness) represents the ratio of true positive results to all actual objects. Their balanced F1-score is expressed as follows:

F_1 = 2 (Precision × Recall) / (Precision + Recall) (5)

To ensure the reliability of the channel estimates, the calibration curve, false positive detections, and stability across various decision thresholds are also examined. Explainable AI maps can reveal whether a detector has focused on the object itself, background features, or arbitrary artifacts.

In an experiment involving UAVs equipped with RGB and thermal sensors for human detection, a customized single-stage detector improved the mean Average Precision (mAP50–95)—calculated across intersection-over-union thresholds ranging from 50% to 95%—by 22% compared to the baseline model. While this result cannot be directly extrapolated to the camouflage of military objects, it demonstrates the value of fusing multimodal information. Furthermore, the proposal of a dual-channel attention network for classifying camouflaged materials using mid-wave infrared hyperspectral imagery points toward a potential direction for future expansion.

Model Limitations

The proposed integrated score does not automatically eliminate detector errors. The R_d value is sensitive to image content, object class, sensor calibration, and the decision threshold. Randomly distributing consecutive frames of the same object across different subsets (training, validation, or testing) leads to data leakage and an artificial inflation of

results. Therefore, data partitioning must be performed based on specific flight missions, geographical areas, or dates of image acquisition.

Another limitation concerns the interpretation of weights. Fixed weights may fail to accurately reflect channel quality under varying weather conditions or times of day; conversely, learnable (trainable) weights carry the risk of overfitting to a small dataset. In practical evaluation, weights—whether fixed, quality-adaptive, or learned—must be compared separately, and confidence intervals should be assessed using resampling methods.

Conclusion

This article proposes the SIMNB model for evaluating the effectiveness of multimodal camouflage for military objects against the visible and thermal surveillance sensors of an unmanned aerial vehicle (UAV). The model integrates detection scores from the visible and thermal channels, along with contextual distinctiveness, into a unified risk metric (R_d) using weights derived from sensor quality. The camouflage criterion provides a relative comparison of risk levels before and after the application of camouflage. The model enables the comparison of different camouflage configurations under identical data and detector conditions, while also identifying the weaker (more vulnerable) modal channel. The cited value of 63.9% is the result of a sample calculation rather than an actual experimental outcome. To validate the model's practical applicability, further requirements include independent RGB and TIR test datasets, sensor calibration, evaluations segmented by flight conditions, and the determination of statistical confidence intervals.

LITERATURE

1. Hua W., Chen Q. A survey of small object detection based on deep learning in aerial images. Artificial Intelligence Review. 2025. 58:162. DOI: 10.1007/s10462-025-11150-9.
2. Hupel T., Stütz P. Adopting hyperspectral anomaly detection for near real-time camouflage detection in multispectral imagery. Remote Sensing. 2022. 14(15):3755. DOI: 10.3390/rs14153755.
3. Eckel L., Stütz P. Hyperspectral sensor management for UAS: performance analysis of context-based system architectures for camouflage and UXO anomaly detection workflows. Drones. 2024. 8(10):529. DOI: 10.3390/drones8100529.
4. Zhou C., Cheng P., Fang J. et al. Optimizing multispectral object detection: a bag of tricks and comprehensive benchmarks. arXiv:2411.18288. 2024.
6. Joshi G., Joshi A., Shetty M. et al. Ensemble learning and EigenCAM-based feature analysis for improving the performance and explainability of object detection in drone imagery. Discover Applied Sciences. 2025. 7:376. DOI: 10.1007/s42452-025-06879-5.
7. Liu M., Hu H., Ma L., Yu Y., Zhang W. COFNet: contrastive object-aware fusion using box-level masks for multispectral object detection. IEEE Transactions on Multimedia. 2025. 27:7444–7458. DOI: 10.1109/TMM.2025.3599097.
8. Sridhar A., Su T.-J. Multispectral human presence detection using adapted YOLO network. 33rd European Signal Processing Conference. 2025. DOI: 10.23919/EUSIPCO63237.2025.11226498.
9. Zhang S., Cao Y., Bai L., Wu Z. Research on camouflage target classification and recognition based on mid-wave infrared hyperspectral imaging. Remote Sensing. 2025. 17(8):1475. DOI: 10.3390/rs17081475.