Choudhary Usman Alam
Instructor : Sir Murtaza taj
Saadi Humayun
TA: Aamil Khan

Project

Wildlife Monitoring : Detection and Classification of Snow Leopards

PAPER 1: Wildlife surveillance using deep learning methods

This paper presents a deep learning-based approach to automate wildlife monitoring. It focuses on species-specific classification using camera-trap images and videos. It uses CNNs for automatic image recognition and classification of animals, particularly badgers.

Relevance to our project:

- The classification approach can be adapted to improve snow leopard detection in our Early Warning System.
- The study proves that deep learning is effective for wildlife monitoring in both day and night scenarios.
- The use of CNNs for image-based classification directly aligns with our goal of enhancing snow leopard recognition.
- The model's success in species identification can help in r alerts.

Key contributions:

- Developed a CNN-based classification model for identifying and differentiating wildlife species.
- Achieved 98.05% accuracy for binary classification (badger vs. other species) and 90.32% for multi-class classification (six species).
- First known application of deep learning for detecting species in video footage in addition to still images.
- Demonstrated the feasibility of AI in reducing human effort for wildlife conservation tasks.

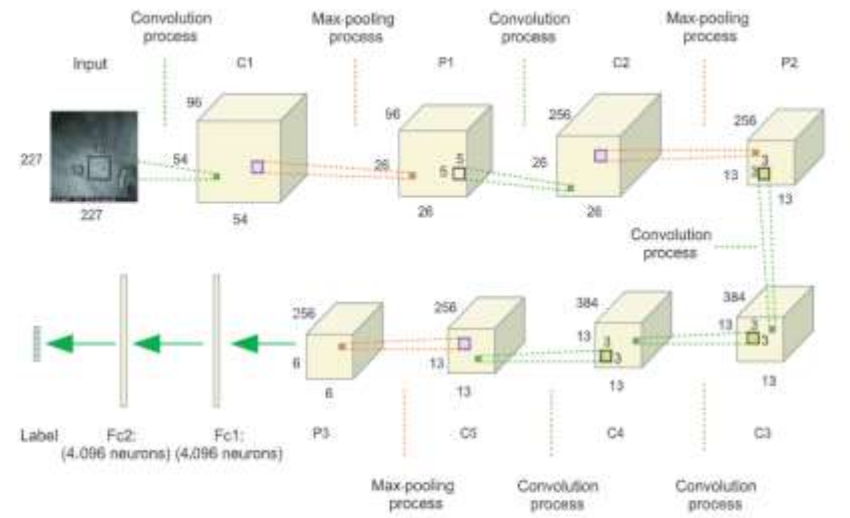


FIGURE 4 The architecture of CNN-2

PAPER 2: Animal Species Recognition Using Deep Learning

This paper presents a deep learning approach using Convolutional Neural Networks to tackle the problem of accurate and efficient animal species recognition from images. The innovation lies in leveraging advanced CNN architectures and training methodologies to achieve improved performance compared to traditional methods, thereby contributing to applications in wildlife monitoring and conservation.

Relevance to our project:

- Paper establishes deep learning (CNNs) as a core technique for animal species recognition, directly relevant to our project.
- Paper discusses challenges like varied lighting and pose, which are key issues in camera trap images.
- Paper covers data augmentation and model optimization, techniques which we will also need to employ for improved accuracy and better generalizability
- Paper's broader goal of using AI for wildlife monitoring mirrors our project's aim to aid snow leopard conservation.

Key contributions:

- The models developed in the paper achieved a high accuracy of 99.8% in detecting objects while 97.6% in identifying the animal’s species.
- Employs two CNNs (cascading filtering) for hierarchical object (human/animal) and species identification to improve efficiency.
- Demonstrated the feasibility of AI in reducing human effort for wildlife conservation tasks.

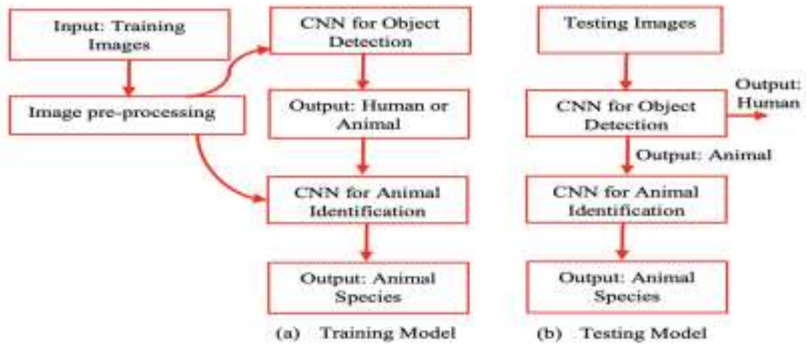


Fig. 3. Recognition system

PAPER 3: Human expertise combined with artificial intelligence improves performance of snow leopard camera trap studies

This study evaluated the use of Whiskerbook AI software to aid in snow leopard identification from camera trap images. It compared the performance of observers with different experience levels and found that AI assistance improved accuracy, but expert observers still provided the most reliable estimates. The findings highlight the potential of AI to reduce errors in wildlife studies that use camera trap data.

Relevance to our project:

- Our project and the paper both focus on the challenge of accurately identifying individual snow leopards from camera trap images.
- Paper's discussion of algorithms like PIE, which use CNNs to extract patterns from snow leopard images can be helpful to our project
- Our plan to use data augmentation and domain adaptation relates to the paper's emphasis on the challenges of image variability (different times, lighting, angles) and the need for robust identification methods.

Key contributions:

- Evaluated Whiskerbook AI software to reduce errors in snow leopard individual identification from camera trap images.
- Compared the performance of "expert" vs. "novice" observers in identifying individual snow leopards with and without AI assistance.
- Showed AI assistance reduced misclassification errors, improving abundance estimates.
- The study utilized the Pose Invariant Embeddings (PIE) algorithm to extract patterns from snow leopard images for individual identification.

Figure 3. Examples of the 8 viewpoints of ObjectPI, for 2 objects.

PAPER 4: Automated Species Identification in Camera Trap Images

This paper introduces an improved ScSPM framework, which crucially combines dense SIFT and cell-structured LBP for local feature extraction and uses weighted sparse coding for dictionary learning. This shows us the practical and effective use of computer vision techniques for wildlife monitoring in complex real-world conditions

Relevance to our project:

- Both this paper and our project directly address the challenge of animal classification from camera trap images
- The paper details the process of building and balancing their dataset from two different field sites and the challenges of dealing with unbalanced initial data
- The paper emphasises the crucial role of local feature extraction for accurate classification which will be useful in our case too.
- The paper provides a detailed overview of the inherent difficulties associated with camera trap images (eg poor and variable illumination etc)

Key contributions:

- A key contribution is the combination of dense SIFT and cLBP to represent the animals in the images
- A linear SVM algorithm is used for the classification of the images
- The paper achieved an average classification accuracy of 82% on the 18-species dataset

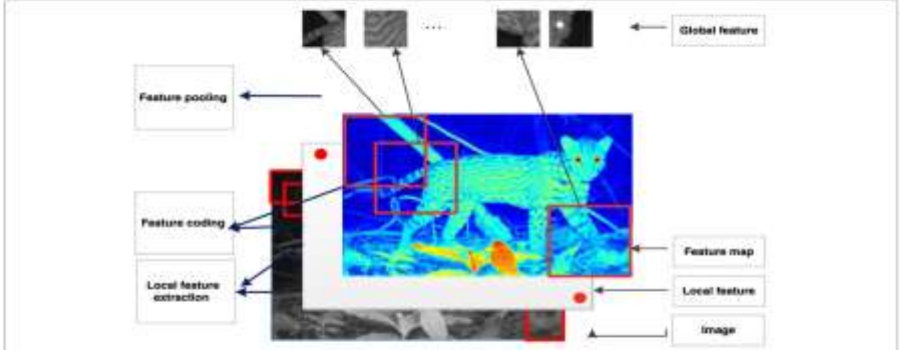


Figure 1 The architecture of ScSPM algorithm. The densely extracted local features are pooled across different spatial locations over different spatial scales.

PAPER 5: YOLO-Unified Real time object detection

This paper provides novel approach to object detection in images. Unlike previous algorithms like R-CNN which used to work in 2 stages, YOLO frames object detection as a single regression problem, predicting bounding boxes and class probabilities in one go using CNN. This simplified pipeline enables remarkably faster processing speeds and helps to generalize to different image domains better than previous models

Relevance to our project:

- Considering YOLO's ability to learn generalizable object representations by seeing the entire image context, as highlighted in its paper regarding artwork generalization, analysing these properties can inform our project's strategies using data augmentation, domain adaptation, and self-supervised learning to improve robustness in camera trap images across varying environmental conditions like night and bad weather.
- Understanding YOLO's error profile, particularly its global reasoning leading to fewer background errors compared to Fast R-CNN, can inspire techniques for better differentiating snow leopards from their background in challenging conditions for our classification task.

Key contributions:

- YOLO significantly simplifies the object detection pipeline compared to previous methods that involved complex, disjoint stages like region proposal, feature extraction, classification, and post-processing
- Unlike sliding window and region proposal-based techniques that only process local regions, YOLO sees the entire image during training and testing
- While YOLO lags behind the most accurate state-of-the-art detection systems, it achieves a significantly higher mAP than other real-time systems

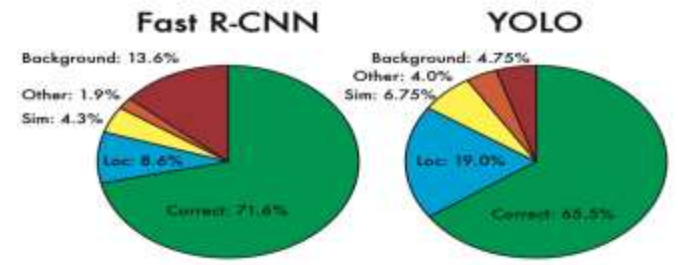


Figure 4: Error Analysis: Fast R-CNN vs. YOLO These charts show the percentage of localization and background errors in the top N detections for various categories (N = # objects in that category).

PAPER 6: Night time wildlife object detection based on YOLOv8-night

This paper is about night time wildlife object detection using infrared cameras, and it introduces a new model called YOLOv8-night1 . This model enhances the YOLOv8 framework by integrating a novel channel attention mechanism designed to improve the accuracy of detecting animals in low-light and complex nocturnal environments.

Relevance to our project:

- Addresses the challenges of detecting wildlife in difficult conditions which can help in our detection method, like in night.
- Emphasises the importance of automated object detection techniques for efficient wildlife monitoring
- Introduces a novel channel attention mechanism that improves the model's ability to focus on animal-related features and suppress noise. We can introduce this channel attention mechanism in our model.
- The comparison with other channel attention mechanisms would help us choose the best for our model.

Key contributions:

- YOLOv8-night, a new object detection model adapted for nighttime wildlife monitoring using infrared cameras
- integration of a new channel attention mechanism into the YOLOv8 framework
- The YOLOv8-night model was tested on real-world nighttime images (NTLNP dataset) and demonstrated excellent performance
- The study provides insights into the optimal placement of the channel attention mechanism within the YOLO architecture.
- The proposed channel attention mechanism was compared with other existing channel attention methods (SENet, ECANet, and CBAM) and was found to outperform them

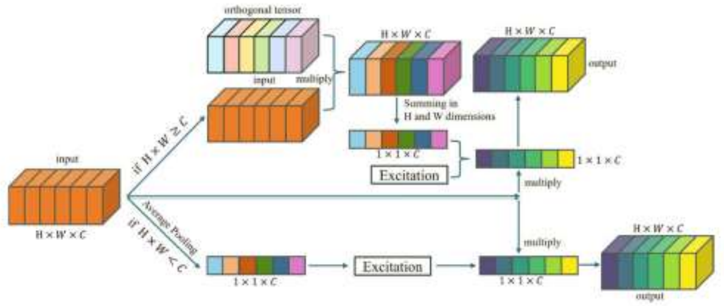


Fig. 1 *Architecture of the channel attention mechanism module. The module comprises two branches: the orthogonal channel compression (OCC) branch and the average pooling channel compression (APCC) branch. When the input feature dimension satisfies $C \leq H \times W$, it will enter the OCC branch, otherwise it will enter the APCC branch, and finally the channel attention weights are obtained by the calculation of the respective branch*

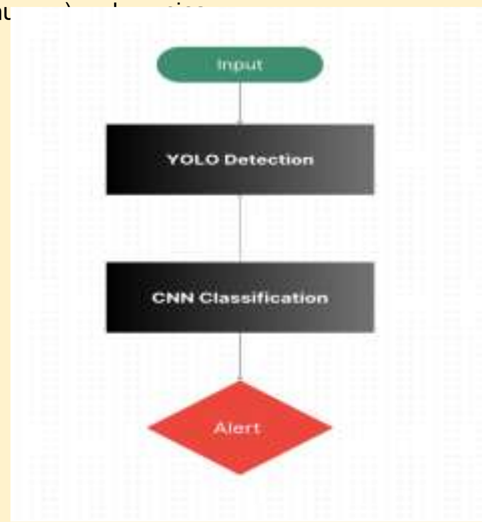
Scope of our Project

Objective

- Improve detection and classification accuracy of snow leopards in camera trap images under challenging conditions (nighttime, bad weather, varied poses).
- Make our model robust enough to function properly under challenging conditions (nighttime, bad weather, varied poses).
- Reduce human-wildlife conflict by enabling real-time alerts

Technical Approach

1. **Model Architecture**
    - Integrate **YOLOv8-night** (Paper 6) for real-time detection in low-light conditions, leveraging its channel attention mechanism.
    - Use **CNNs** (Papers 1, 2, 3) for species-specific classification, inspired by cascading filtering (Paper 2) and Pose Invariant Embeddings (Paper 3).
    - Explore hybrid architectures combining YOLO's speed with CNNs' accuracy for hierarchical detection (animal vs. hu identification.
2. **Robustness Enhancements**
    - **Data Augmentation**: Simulate snow, fog, and lighting variations (Paper 2, 4).
    - **Domain Adaptation**: Transfer learning to generalize across seasons and camera locations (Paper 3, 4).
    - **Self-Supervised Learning**: Pretrain on unlabeled data to mitigate limited labeled datasets (Paper 5).



Data Handling

- Curate a balanced dataset of snow leopards, domestic animals, and humans from camera traps.
- Address class imbalance using techniques from Paper 4 (ScSPM framework) and synthetic data generation

Limitations & Boundaries

- Focus on snow leopards and common domestic animals (goats, sheep).
- Prioritize software/model development over hardware modifications.