From CNN-Based Indoor Scene Understanding to Automatic Inference of Burnt Object Remnants: A New Research Frontier

[1] *Ka-Fai Felix Hung* (洪嘉徽), [1] *Chiou-Shann Fuh* (傅楸善)

[1]Department of Computer Science and Information Engineering,
National Taiwan University, Taipei, Taiwan,

[*]E-mail: b09902090@csie.ntu.edu.tw fuh@csie.ntu.edu.tw

Abstract

Post-fire environments pose a significant challenge for forensic investigators and recovery teams. Physical objects are often reduced to heavily damaged or fragmented debris, making it difficult to determine their original identities. Accurate reconstruction of pre-fire object configurations from such debris is critical for fire investigation, insurance assessment, and post-disaster recovery planning.

This study proposes a conceptual framework that uses Convolutional Neural Networks (CNNs) to analyze images of fire-damaged scenes and infer the identities of the original objects. The problem is formulated as an inverse reconstruction task, where the goal is to classify object remnants and approximate their pre-fire states based on incomplete visual cues. The proposed system integrates CNN-based object fragment recognition with spatial reasoning to enhance inference accuracy.

Due to the scarcity of real-world datasets, synthetic burned-debris images are generated using controlled data augmentation techniques to simulate post-fire visual characteristics. While the present work focuses on system design and conceptual validation, it outlines a practical direction for future implementation and testing. This approach is intended to support automated fire forensics, aid recovery operations, and improve the efficiency of post-fire scene analysis.

Keywords: *Burned Object Inference, Post-Fire Scene Analysis, Deep Learning, CNN, Forensic Reconstruction, Synthetic Data Augmentation*

1. INTRODUCTION

Fires remain among the most destructive forms of disaster, often leaving behind environments filled with

damaged and heavily altered debris. For investigators, recovery personnel, and insurance professionals, it is essential to determine what objects were present in a scene prior to the fire. Identifying these objects from burned remnants is a difficult task, as visual evidence is frequently incomplete and distorted by heat, soot, and structural collapse.

Traditionally, post-fire object identification has relied on manual inspection by human experts. This process requires significant time and domain-specific knowledge, and is often hindered by the subjective nature of visual interpretation in complex, degraded environments. The growing availability of digital imagery from fire scenes, captured using drones, mobile devices, or inspection robots, creates an opportunity for automated analysis to assist human investigators.

Recent advances in deep learning, particularly Convolutional Neural Networks (CNNs), has shown promise in visual recognition tasks across many domains. However, applying these techniques to post-fire scenes introduces unique challenges. Object appearances are radically altered by fire damage, and intact datasets of burned objects are scarce. Moreover, the task is not simply to recognize objects but to infer their pre-fire identities from partial and degraded information.

Our study proposes a conceptual framework for post-fire object inference using CNN-based visual analysis. The system aims to detect and classify object fragments in burned scenes, and to reason about their original identities and spatial configurations. To address the lack of real-world training data, the framework incorporates synthetic data generation methods that simulate post-fire visual characteristics on known objects. Spatial modeling is also introduced to help interpret ambiguous fragment information by considering object relationships and scene context.

The contributions of this work include:

- A conceptual system design for inferring original object identities from fire-damaged debris using CNNs.

- A synthetic data generation approach that enables training and evaluation in the absence of large real-world post-fire datasets.

- A spatial reasoning component that models object relationships to improve inference accuracy in highly degraded scenes.

- A discussion of potential applications, including automated fire forensics, insurance assessment, and post-disaster recovery support.

We focuses on establishing a theoretical foundation and system architecture for future research. The proposed framework offers a promising direction for integrating artificial intelligence into post-fire analysis workflows, with the goal of improving accuracy, efficiency, and consistency in this challenging domain.

2. Background and Objectives

Background

Post-fire environments present extreme challenges for visual interpretation. Furniture, appliances, and other indoor items are often subjected to intense heat, combustion, and structural collapse. As a result, their physical forms are altered beyond easy recognition. Materials may melt, warp, or disintegrate, leaving only charred remnants or fragments partially embedded in surrounding debris.

In many practical scenarios, such as insurance investigations, building recovery planning, and fire origin analysis, it is crucial to determine what objects existed in the scene before the fire. This includes identifying high-value items, locating potential fire sources, or reconstructing room functions based on former object presence. Currently, these tasks are performed manually by experienced fire investigators or adjusters who rely on their domain knowledge and past case references. However, such manual methods are limited by time, subjectivity, and inconsistent documentation.

Advances in computer vision, especially through deep learning and Convolutional Neural Networks (CNNs), offer new possibilities for automating object identification. While CNNs have been widely applied in standard image recognition tasks, their potential for post-disaster applications has not been fully explored. The specific challenges of burned environments（such as extreme deformation, soot occlusion, and incomplete shapes）require a specialized inference approach that combines robust feature extraction with spatial reasoning and prior knowledge.

We aims to establish a conceptual framework for the automated inference of pre-fire object identities based on images of post-fire debris. The framework uses CNNs to recognize surviving object fragments and to estimate their original categories and possible configurations. Synthetic image degradation is introduced to simulate burned object appearance, enabling model training in the absence of real-world labeled data. The system is intended to serve as a tool for experts in fire investigation and recovery, supporting more accurate and timely post-incident assessment.

Objectives

This research seeks to define a viable system architecture and algorithmic pipeline that address the following core objectives:

1.Establish a Damage-Tolerant Object Inference Method
Develop a deep learning approach capable of recognizing object fragments and predicting their original identity despite extensive damage. The system must handle scenarios with high visual distortion, missing parts, and inconsistent lighting.

2. Simulate Fire Damage Using Synthetic Data
Create realistic burned-object images through simulation techniques. This includes applying visual effects such as melting, charring, blurring, occlusion, and partial fragmentation to known datasets of intact objects.

3. Integrate Spatial Reasoning in Scene Reconstruction
Use object location, shape, and contextual clues to

reconstruct likely spatial layouts. Scene graph modeling and rule-based constraints may be incorporated to infer typical arrangements or detect object co-occurrence patterns.

4. Support Practical Post-Fire Applications
Ensure that the proposed system aligns with real-world needs in fire forensics, insurance assessment, and emergency response documentation. The design must support scalable deployment in visually diverse and severely damaged environments.

By addressing these objectives, the framework aims to contribute both to the field of computer vision and to practical applications in disaster recovery and forensic investigation.

3. Related Work

3.1 Object Recognition under Visual Degradation

Object recognition in computer vision has advanced significantly with the introduction of deep learning models. CNN-based architectures such as Deep Residual Networks (ResNet[1]), MobileNet, and EfficientNet have demonstrated strong performance in standard classification and detection tasks. However, most of these achievements are based on clean, unaltered images. In burned environments, objects suffer from irregular damage patterns, non-uniform surfaces, and partial occlusion, which reduce the effectiveness of traditional recognition pipelines.

Related domains, such as defect detection in manufacturing and medical anomaly detection, provide some insight into handling abnormal visual conditions. Models trained on degraded or irregular data often benefit from techniques such as transfer learning, fine-tuning with domain-specific samples, and attention-based feature enhancement. These strategies can inform the design of recognition systems for post-fire scenes.

3.2 Post-Fire Scene Analysis and Reconstruction

Fire scene analysis traditionally focuses on determining the origin and cause of the fire, often using photographs and physical inspection. Some studies have applied photogrammetry or 3D scanning to document fire damage, but there is limited research on automated object recognition in post-fire imagery.

Few existing works tackle the inverse problem of reconstructing what was originally present in a fire-damaged room. One notable challenge is the absence of large-scale datasets of post-fire scenes with annotated object identities. Without sufficient training data, models cannot generalize to the wide variety of damage types and materials found in real incidents.

Nonetheless, recent efforts in disaster informatics and autonomous inspection systems show that intelligent virtual agents can assist in post-event documentation. These trends support the feasibility of applying AI techniques to post-fire object inference, provided that appropriate training strategies are in place.

3.3 Synthetic Data Generation and Domain Adaptation

When labeled real-world data is scarce, synthetic data generation offers a practical alternative. Techniques such as image augmentation[6], generative adversarial networks (GANs[7]), and damage simulation allow researchers to create proxy datasets that mimic real-world conditions. These synthetic datasets can then be used to pre-train models or to evaluate conceptual architectures.

Domain adaptation further enhances model robustness by aligning feature distributions between synthetic and real images. Approaches such as unsupervised transfer learning, cycle-consistent image translation, and domain adversarial training have been applied successfully in other fields, including medical imaging and autonomous driving. Similar methods may be employed to bridge the gap between simulated burned-object data and actual fire scene photos.

3.4 Summary

Although the problem of post-fire object inference remains largely unexplored, several areas of existing research offer useful methodologies and inspiration. Deep learning provides a conceptual approach for extracting patterns from complex visual data. Synthetic data generation and domain adaptation allow researchers to overcome data scarcity. Spatial reasoning and scene graph modeling offer tools to interpret partial evidence in cluttered environments. By integrating these approaches, the proposed framework aims to initiate a new direction in automated post-disaster scene understanding.

4. Methodology

This study proposes a conceptual system for post-fire object inference based on deep learning. The system consists of three main modules: synthetic data preparation, damage-tolerant object recognition, and spatial reasoning for scene reconstruction. Each component is designed to address the specific challenges posed by burned environments.

4.1 Synthetic Data Preparation

Since real-world datasets of burned objects with reliable annotations are not readily available, this framework adopts a synthetic data generation strategy. Existing indoor object image datasets, such as ImageNet Large Scale Visual Recognition Challenge dataset (ImageNet) [4] or COCO, are used as the source of intact object images. Controlled visual degradation techniques are then applied to simulate fire damage effects.

The degradation pipeline includes the following transformations:

- Charring and Soot Simulation: Darkening and texture overlay to mimic soot deposits and burned surfaces.

- Melting and Warping: Geometric distortion applied to object boundaries to simulate heat-induced deformation.

- Occlusion and Fragmentation: Random masking and partial occlusion to simulate object breakage or coverage by debris.

- Blur and Noise: Application of blur and image noise to represent smoke or camera limitations in post-fire conditions.

By systematically applying these effects with varying parameters, a diverse dataset of simulated burned object images is created. These images are paired with the original object labels, enabling supervised learning of damage-tolerant recognition models.

4.2 Damage-Tolerant Object Recognition

The core of the system is a Convolutional Neural Network (CNN) trained to recognize object fragments in burned scenes. Pre-trained models such as ResNet[1] or EfficientNet serve as the base feature extractors, fine-tuned using the synthetic burned dataset.

Key considerations include:

- Robust Feature Extraction: The model must learn to identify invariant features that persist even under significant damage.

- Partial Object Detection: The system must be capable of making category predictions based on incomplete visual cues.

- Multi-Scale Learning: The network architecture may incorporate multi-scale feature layers to improve detection of small or fragmented object parts.

A practical option is to adopt a Faster Region-based Convolutional Neural Networks (Faster R-CNN) [2] or Mask Region-based Convolutional Neural Networks (Mask R-CNN) [3] framework may be adopted to enable both object classification and localization within cluttered fire scene images.

4.3 Spatial Reasoning and Scene Reconstruction

To enhance inference accuracy, the system incorporates spatial reasoning through scene graph construction. After object fragments are detected, their bounding boxes and relative positions are used to build a graph where nodes represent object candidates and edges represent spatial relationships.

The reasoning process involves:

- Object Co-occurrence Modeling: Incorporating prior knowledge of common object pairings and arrangements to resolve ambiguities.

- Layout Estimation: Estimating the probable pre-fire layout of the room based on detected objects and their spatial configuration.

- Rule-Based Refinement: Applying constraints such as expected object sizes, proximity to structural elements (walls, floor), and typical placement patterns.

This combined reasoning supports more plausible reconstruction of the original scene, even when individual object detections are uncertain or incomplete.

4.4 System Integration

The complete system operates as an end-to-end pipeline:

1. Input burned scene images.

2. Detect and classify object fragments using the trained CNN model.

3. Construct a scene graph representing detected object positions and relationships.

4. Infer likely pre-fire object categories and spatial configuration.

5. Output a visual reconstruction and a structured report of inferred objects.

This architecture is designed for extensibility, allowing future integration of additional data sources such as depth maps, thermal imagery, or 3D scans.

5. Discussion and Future Work

5.1 Limitations

The proposed system faces several inherent challenges:

- **Data Realism:** Simulated burned images may not fully capture the diversity and complexity of real fire damage. Domain adaptation will be necessary to bridge this gap.

- **Object Fragment Ambiguity:** Severely damaged fragments may resemble multiple object classes, leading to uncertain predictions.

- **Scene Context Variability:** The diversity of room layouts and object configurations may limit the generalizability of spatial reasoning models.

5.2 Future Directions

Several enhancements are envisioned for future research:

- **Domain Adaptation Techniques:** Employ advanced transfer learning and adversarial domain adaptation to improve model robustness on real post-fire imagery.

- **Multi-Modal Integration:** Incorporate additional modalities such as depth sensing, infrared imaging, or material analysis to provide richer input data.

- **Temporal Analysis:** Use pre-fire floorplans or archival images where available to guide reconstruction.

- **Interactive Human-in-the-Loop Systems:** Enable human experts to interactively validate and refine AI-generated scene reconstructions.

5.3 Practical Applications

Potential applications of this system include:

- **Fire Forensics:** Supporting investigators in identifying possible fire sources and reconstructing room content.

- **Insurance Assessment:** Automating object identification to streamline claims processing and damage valuation.

- **Disaster Recovery:** Assisting architects and planners in restoring damaged spaces.

- **Heritage Preservation:** Reconstructing valuable artifacts or culturally significant spaces after fire damage.

By pursuing these directions, the proposed framework can evolve into a valuable tool for automated post-fire analysis, reducing the manual burden on experts and improving the accuracy and consistency of post-incident investigations.

Future work will also include systematic analysis of module contributions, such as the effects of scene graph reasoning and damage-aware feature learning on overall inference performance.

6. Error Analysis

Although the proposed system is conceptual, several expected sources of error can be anticipated based on the challenges inherent in post-fire visual analysis. Understanding these limitations is critical for guiding future development and improving system robustness.

6.1 Fragment Ambiguity

Severely damaged object fragments may lose key distinguishing features. For example, a charred metallic frame could originate from a chair, a table, or a shelving unit. Without clear contextual clues, CNN models may confuse such fragments, leading to misclassification.

Additionally, certain materials such as plastics may melt into amorphous shapes that provide little useful structure for recognition. Textures, edges, and keypoints commonly used in standard object detection are often absent or unreliable in burned debris.

6.2 Background Clutter

Post-fire environments are typically filled with dense, irregular clutter. Ash, soot, and collapsed structural elements may dominate the scene, introducing high visual noise. Small objects or thin fragments may be lost in this clutter or falsely detected as parts of larger items.

False positives may arise when background patterns or irregular debris resemble object-like features. Conversely, true object fragments may be occluded or blended with surrounding debris, causing false negatives.

6.3 Scale and Perspective Distortion

Burned scenes are often captured under suboptimal conditions. Wide-angle cameras, handheld devices, or drone footage may introduce perspective distortion and inconsistent scale. Objects may appear flattened, stretched, or skewed relative to their original form.

Such distortions complicate both object recognition and spatial reasoning. Bounding box accuracy may suffer, affecting the reliability of scene graph construction and layout inference.

6.4 Over-Reliance on Synthetic Data

While synthetic burned-debris data provides a necessary starting point for model training, it may not fully capture the variability of real-world fire damage. Real burned scenes exhibit diverse material behaviors, lighting conditions, and post-event interventions (such as firefighting water damage).

Models trained solely on synthetic data may generalize poorly without effective domain adaptation. Unexpected visual artifacts in real images may trigger classification errors or cause the system to miss key object fragments.

7. Dataset Description and Generation

Given the scarcity of publicly available post-fire object datasets, we proposes a synthetic data generation approach to support model development and evaluation.

7.1 Source Datasets

Existing large-scale object recognition datasets serve as the source of intact object images. Suitable candidates include:

- **ImageNet[4]:** A diverse collection of object images across thousands of categories.

- **COCO (Common Objects in Context):** Provides object annotations within complex scenes, supporting realistic context-aware degradation.

- **SUN RGB-D:** Offers indoor scene images with rich object annotations and depth information, useful for simulating spatial context.

7.2 Synthetic Degradation Pipeline

To simulate post-fire appearance, the following visual degradation techniques are applied:

- **Color Transformation:** Grayscale conversion, localized darkening, and color jitter to mimic soot deposits and heat discoloration.

- **Texture Overlay:** Application of burned textures and noise patterns to replicate surface charring.

- **Geometric Distortion:** Mesh warping and perspective shifts to simulate melting, bending, or collapse.

- **Occlusion:** Random partial masking to represent fragmentation or debris coverage.

- **Blur and Noise:** Addition of Gaussian blur and noise to replicate smoke or imaging artifacts.

The degradation parameters are randomized within realistic ranges to generate diverse training samples. Each degraded image retains its original object label, supporting supervised learning.

7.3 Dataset Split and Usage

The synthetic burned dataset is divided into training, validation, and test subsets. During system development, models are trained on the synthetic data and validated using a holdout set. Qualitative validation may also be performed on selected real post-fire images, where available.

Future work will involve collecting and annotating real-world post-fire imagery to supplement synthetic data and support more rigorous evaluation.

8. Conclusion

Our study proposes a conceptual framework for automated post-fire object inference using deep learning. By leveraging Convolutional Neural Networks, synthetic data generation, and spatial reasoning, the system aims to infer the identities and configurations of objects in burned environments.

Key contributions include:

- Defining a damage-tolerant object recognition pipeline capable of handling heavily degraded visual inputs.

- Designing a synthetic data generation process to address the lack of real-world burned object datasets.

- Integrating spatial reasoning through scene graph modeling to improve inference accuracy and scene-level consistency.

- Outlining potential applications in fire forensics, insurance assessment, disaster recovery, and cultural heritage preservation.

While this work focuses on system design and conceptual validation, it establishes a foundation for future research and practical implementation. The challenges of fragment ambiguity, cluttered backgrounds, and domain adaptation must be addressed through continued refinement and the incorporation of real-world data.

This work is intended to provide a foundation for future tools that could assist human experts in performing faster, more accurate, and more consistent assessments of fire-damaged environments.

9. Qualitative Results

To illustrate the potential capabilities of the proposed system, this section presents conceptual examples of its expected outputs. While a fully trained model is not yet implemented, qualitative analysis based on synthetic data and prototype visualizations offers insights into practical applications.

9.1 Example Scenarios

Insurance Assessment of Residential Fire

A post-fire image of a living room is processed by the system. Despite extensive charring and structural collapse, the system detects fragmentary remains of a sofa frame, a melted television base, and the metallic components of a floor lamp.

The reconstructed scene graph identifies these objects and infers their likely pre-fire positions. An automated report lists:

- **Detected objects:** Sofa (high confidence), Television (medium confidence), Floor lamp (high confidence)

- **Spatial arrangement:** Sofa against east wall, television on floor near west wall, lamp adjacent to sofa.

This output aids insurance adjusters in documenting the scene and validating claims regarding lost property.

Fire Forensics Investigation

In an office building fire, the system analyzes images from a drone survey. It detects fragments of a desk, a set of filing cabinets, and remnants of electrical equipment near the point of origin.

By correlating object types and spatial proximity, investigators are supported in identifying potential ignition sources and reconstructing the original room function.

Cultural Heritage Restoration

A fire damages part of a historic library. The system processes post-event imagery and identifies metallic fragments corresponding to archival shelving and charred remains of wooden reading desks.

This information assists restoration experts in planning reconstruction, prioritizing the recovery of historically significant structural elements.

9.2 Visual Concept Examples

Prototype visualizations demonstrate the following:

- **Object fragment detection:** Bounding boxes around detected remnants, labeled with predicted object categories.

- Scene graph overlay: Visual graph showing inferred relationships and spatial arrangement of detected objects.

- Original layout approximation: Suggested reconstruction of the pre-fire room configuration based on object positions.

These qualitative outputs support the envisioned role of the system as an assistive tool for human experts.

10. Comparative Visualization and Segmentation-Based Enhancement

To conceptually evaluate the added value of segmentation, the system may be extended with a Mask R-CNN[3] framework, enabling pixel-level delineation of object fragments.

10.1 Expected Benefits

- Improved object localization: Fine-grained masks help separate overlapping fragments.

- Enhanced spatial reasoning: More precise shape information supports better scene graph construction.

- Support for robotic applications: Accurate segmentation assists robots in identifying and manipulating debris during recovery operations.

10.2 Visual Comparison

Conceptual comparison between bounding box detection and segmentation-based outputs shows:

- Bounding box only: Fragments partially enclosed, potential overlap between objects.

- Segmentation masks: Clearer separation of individual object remnants, improved interpretability.

In practice, segmentation is expected to improve recognition performance in highly cluttered and fragmented post-fire environments.

While quantitative comparisons are beyond the scope of this conceptual study, future work will include benchmarking against baseline models once real or large-scale synthetic datasets become available.

References

[1]He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *CVPR*.

[2]Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *NeurIPS*.

[3]He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. *ICCV*.

[4]Russakovsky, O., Deng, J., Su, H., et al. (2015). ImageNet Large Scale Visual Recognition Challenge. *IJCV*.

[5]Krishna, R., Zhu, Y., Groth, O., et al. (2017). Visual Genome: Connecting Language and Vision Using Crowdsourced Dense Image Annotations. *IJCV*.

[6]Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*.

[7]Goodfellow, I., Pouget-Abadie, J., Mirza, M., et al. (2014). Generative Adversarial Nets. *NeurIPS*.

[8]Zhang, Y., Murez, Z., Wang, T., et al. (2018). Deep Contextualized Scene Understanding for Robotics. *RA-L*.