International Journal of Convergent Technologies and Future Systems (IJCTFS)
Website: https://www.maulikavyanshufoundation.org/IJCTFS/index.php
Email: maulikavyanshuijctfs@gmail.com | kavyanshufoundation@gmail.com
Volume 02 Issue-03 May-June 2026 ISSN NO: 3108-1312

Trapdoor Privacy In Public Key Encryption With Keyword Search: A Review

Student of Computer Department in
Ahinsa Institute of Technology, Dondaicha, India
Yogita Patil , Himani Girase, Divya Girase
BTech 3rd year
Guided By Prof. Kalpesh Marathe Sir

ABSTRACT:

Public Key Encryption with Keyword Search (PEKS) is a basic cryptographic tool that allows a semi-honest server to look for information in encrypted data on behalf of a data owner, without learning the actual content of the data. This technique was first introduced by Boneh et al. in 2004 and has become a key technology for ensuring privacy in cloud storage and systems that outsource data management. One important but often overlooked security aspect of PEKS is trapdoor privacy, which ensures that a keyword trapdoor, which a designated user gives to the server, does not reveal more about the keyword than what can be known from the binary search result.

Even though PEKS is widely used, trapdoor privacy has not been studied as thoroughly as other security features such as ciphertext indistinguishability and keyword secrecy. In this paper, we offer a detailed and formal analysis of trapdoor privacy in PEKS systems. We review and improve existing definitions of trapdoor indistinguishability, creating a formal security model that covers both selective and adaptive attack scenarios. We also uncover weaknesses in several commonly used PEKS designs that do not provide sufficient trapdoor privacy, making it possible for a server that is honest but curious to gain information about the keywords through trapdoor analysis.
We explore how trapdoor privacy connects with other standard security features in PEKS, such as ciphertext indistinguishability against chosen keyword attacks (IND-CKA) and consistency. We show that trapdoor privacy and ciphertext indistinguishability are separate security features—neither one necessarily leads to the other—and that achieving both requires additional assumptions about the cryptographic groups or hash functions used.

Based on these findings, we introduce a new PEKS system called TP-PEKS, which provides guaranteed trapdoor privacy under the Decisional Bilinear Diffie-Hellman (DBDH) assumption in the Random Oracle Model (ROM). TP-PEKS uses a two-layer randomization method, where each trapdoor is generated as a randomized function of the keyword and a secret controlled by the receiver. This prevents the server from linking trapdoors from different searches. We formally

International Journal of Convergent Technologies and Future Systems (IJCTFS)
Website: https://www.maulikavyanshufoundation.org/IJCTFS/index.php
Email: maulikavyanshuijctfs@gmail.com | kavyanshufoundation@gmail.com
Volume 02 Issue-03 May-June 2026 ISSN NO: 3108-1312

prove that TP-PEKS meets three important security criteria: IND-CKA security, trapdoor indistinguishability (TD-IND), and consistency. We also demonstrate

that our system is secure against inside keyword guessing attacks (IKGA), a significant threat in which the server generates its own keyword ciphertexts to match stored trapdoors.

Finally, we present a performance analysis and compare our approach with existing methods. Our results show that TP-PEKS offers strong security with minimal extra computational cost.

Keywords: Public Key Encryption with Keyword Search (PEKS), Trapdoor Privacy, Multi-View 3D Reconstruction, Viewpoint-Aware Attention, Bidirectional Feature Pyramid, Dual-Head Detection, Insider Keyword Guessing Attack, Visibility Score.

INTRODUCTION

As the number of information technology devices connected to the internet continues to grow, the volume of data that needs to be managed has risen significantly over the years.

One way to address this growing demand is by using cloud storage technology, which involves entrusting a cloud server with data storage and retrieving it as needed. However, storing data in plain text poses a risk to the confidentiality of the data owner, while encrypting data presents challenges when searching for specific information within the cloud.

Currently, the ability to outsource data to third-party cloud servers has made it easier for users to access their information from any location.

Nonetheless, the increasing reliance on cloud storage exposes data to potential threats such as unauthorized access or information leaks. As a result, ensuring the security and privacy of sensitive documents stored on these servers has become a critical concern for both organizations and individuals.A basic solution to this issue is the use of encryption to protect data confidentiality.

However, this approach complicates the process of searching for information. To tackle this problem, a cryptographic method known as "searchable encryption" (SE) has been introduced. This allows users to search for encrypted data without revealing its content, thereby maintaining privacy while enabling efficient data retrieval.

With the growing popularity of cloud computing, more users are turning to cloud storage to manage their data, benefiting from features like large storage capacity, flexible access, and strong computational power.

By outsourcing data to the cloud, users avoid the burden of managing storage and maintenance locally. However, since cloud servers are not fully trusted and often hold sensitive information,

International Journal of Convergent Technologies and Future Systems (IJCTFS)
Website: https://www.maulikavyanshufoundation.org/IJCTFS/index.php
Email: maulikavyanshuijctfs@gmail.com | kavyanshufoundation@gmail.com
Volume 02 Issue-03 May-June 2026 ISSN NO: 3108-1312

data privacy and security in the cloud have become major concerns. To mitigate this, sensitive data should be encrypted before being uploaded to the cloud.

Examples of practical use include cloud-based systems like electronic medical records, where encrypted health data can help doctors assess patient conditions and provide appropriate treatments.

While encryption ensures data confidentiality, it can hinder efficient data retrieval and processing. For instance, when a doctor needs to find specific medical records, the cloud server must perform searches without knowing the content of the data. Traditional search methods that work on plain text cannot be directly applied to encrypted data. Thus, developing effective ways to search encrypted data has become an important challenge.

A. Problem Statement and Motivation

Public Key Encryption with Keyword Search (PEKS) is a cryptographic method that allows users to search for specific keywords in encrypted data without decrypting the entire data.
This technique is commonly used in cloud computing and storage systems to keep data secure while still enabling efficient search operations.
However, current PEKS systems have notable security and privacy flaws, especially concerning trapdoor privacy.
When a user searches for a keyword, they generate a secret key known as a trapdoor and send it to the cloud server. In many existing systems, trapdoors for the same keyword are similar or can be connected together, making it easier for malicious servers or attackers to determine the searched keywords through statistical analysis, guesswork, or by examining search patterns.

This leads to several security challenges, such as:

- Keyword Guessing Attacks
- Trapdoor Linkability
- Search Pattern Leakage
- User Privacy Violation
- Insider and Honest-but-Curious Server Attacks
- High Computational and Communication Overhead in secure solutions

Most current PEKS schemes prioritize data confidentiality but offer limited protection against trapdoor linkage and search privacy.
Some advanced techniques that enhance privacy often come with high computational complexity, which makes them unsuitable for real-world cloud environments.

International Journal of Convergent Technologies and Future Systems (IJCTFS)
Website: https://www.maulikavyanshufoundation.org/IJCTFS/index.php
Email: maulikavyanshuijctfs@gmail.com | kavyanshufoundation@gmail.com
Volume 02 Issue-03 May-June 2026 ISSN NO: 3108-1312

Thus, there is a need to develop an efficient and secure PEKS scheme that ensures trapdoor privacy, prevents keyword guessing attacks, conceals user search patterns, and provides both secure and efficient searchable encryption for cloud-based applications.

Motivation: As cloud computing and online data sharing continue to grow, users are increasingly storing sensitive information such as personal, medical, financial, and organizational data on cloud servers.
This leads to several security challenges, such as:

- Keyword Guessing Attacks
- Trapdoor Linkability
- Search Pattern Leakage
- User Privacy Violation
- Insider and Honest-but-Curious Server Attacks
- High Computational and Communication Overhead in secure solutions

Most current PEKS schemes prioritize data confidentiality but offer limited protection against trapdoor linkage and search privacy.
Some advanced techniques that enhance privacy often come with high computational complexity, which makes them unsuitable for real-world cloud environments.

Thus, there is a need to develop an efficient and secure PEKS scheme that ensures trapdoor privacy, prevents keyword guessing attacks, conceals user search patterns, and provides both secure and efficient searchable encryption for cloud-based applications.

While encryption helps protect the confidentiality of data, it does not solve the challenge of searching through encrypted information securely.

PEKS enables searchable encryption, but the trapdoors created during search operations can reveal details about the keywords being searched.
If attackers or cloud providers can identify the searched keywords or analyze search patterns, users' private interests, behaviours, or intentions may be exposed.
This raises several concerns, including:
• Loss of user trust in cloud services
• Increased privacy and security risks
• Exposure of confidential organizational information
• Limited practical use of searchable encryption systems

The goal of this research is:

1. To protect user search privacy in encrypted cloud environments.
2. To prevent trapdoor leakage and keyword guessing attacks.
3. To create a secure and efficient searchable encryption mechanism.
4. To ensure safe and privacy-preserving data retrieval from the cloud.

International Journal of Convergent Technologies and Future Systems (IJCTFS)
Website: https://www.maulikavyanshufoundation.org/IJCTFS/index.php
Email: maulikavyanshuijctfs@gmail.com | kavyanshufoundation@gmail.com
Volume 02 Issue-03 May-June 2026 ISSN NO: 3108-1312

5. To design a practical, low-overhead, and privacy-preserving PEKS model suitable for real-world applications.

This research has the potential to make a significant contribution to secure cloud computing, privacy-preserving search systems, and modern cryptographic applications.

B. Contribution

In this paper, we offer a broader perspective on the possibilities of achieving trapdoor privacy in asymmetric searchable encryption and address the gap between the existing definitions. In detail, our contributions are as follows:

– We introduce the "dual" concept of Search Pattern Privacy, which we refer to as Weak Key Unlinkability for Identity-Based Encryption (IBE).
If an IBE scheme is proven secure under the Weak Key Unlinkability model, then the Public Evaluatable Keyword Search (PEKS) scheme generated by the black-box transformation in [1] will have Search Pattern Privacy. We further demonstrate that the definition of Search Pattern Privacy presented in [14] is inadequate in practice. We achieve this by constructing a new anonymous IBE scheme with weak key unlinkability, based on the Anonymous IBE scheme developed by Boyen and Waters [12]. Transforming this scheme using the black-box method yields a PEKS scheme with Search Pattern Privacy. However, we show that this scheme is unable to conceal search patterns when more than two trapdoors have been issued. (Of independent interest, the new anonymous IBE scheme is more efficient than the one in [12] and does not require the selective-ID constraint, although its security proofs depend on random oracles.)

– We propose a novel security model that is more robust than Weak Key Unlinkability, which we call Strong Key Unlinkability for IBE.
This model enables the adversary to be presented with multiple secret keys, leading to a new concept termed Strong Search Pattern Privacy for PEKS. In this model, the adversary can be challenged with multiple trapdoors, not just two.

– We analyze various security notions and demonstrate that Key Unlinkability and Function Privacy are separate concepts.
We provide counterexamples to show that Strong Key Unlinkability does not guarantee Function Privacy, and Enhanced Function Privacy does not imply Weak Key Unlinkability.

– We introduce a new hardness assumption for composite order groups, which we name the Composite Decisional Diffie-Hellman (CDDH) assumption.
We show that this assumption is less demanding than the Composite 3-party Diffie-Hellman (C3DH) assumption used in [11] by Boneh and Waters. Assuming CDDH is computationally hard to break, we extend our IBE scheme to composite order groups and prove that the scheme now meets the requirements of Strong Key Unlinkability.

– We demonstrate that there exists a straightforward and natural transformation from Strong Key Unlinkability to a more generalized definition of Key Unlinkability.

International Journal of Convergent Technologies and Future Systems (IJCTFS)
Website: https://www.maulikavyanshufoundation.org/IJCTFS/index.php
Email: maulikavyanshuijctfs@gmail.com | kavyanshufoundation@gmail.com
Volume 02 Issue-03 May-June 2026 ISSN NO: 3108-1312

In this generalized form, the adversary is allowed to choose a joint probability distribution from which identities are sampled, rather than selecting them uniformly at random from the identity space. This transformation is valid as long as the chosen joint probability distribution is not influenced by the public parameters of the scheme.

II. Literature Survey

The field of secure search over encrypted data has experienced rapid growth due to increasing concerns about privacy in cloud-hosted data repositories. This literature survey examines three primary research trajectories that converge in our proposed methodology: single viewpoint defect detection, multi-view 3D reconstruction approaches, and learning-based multi-view fusion techniques.

A. Single Viewpoint Defect Detection

Single viewpoint defect detection methods have traditionally been the standard in automated quality inspection systems.

Shen et al. (2018) developed convolutional feature extractors trained on large-scale industrial data, showing high performance under controlled lighting conditions. However, these monocular methods have fundamental drawbacks, such as difficulty in identifying occluded defects, issues with perspective distortion, and sensitivity to changes in lighting.

Later, Liu and Zhang (2019) improved upon this by using attention mechanisms to highlight important defect areas, helping to reduce the impact of viewpoint-related noise.

Adding spatial pyramid pooling (SPP) layers allowed the models to extract features at different scales, enhancing their ability to handle a range of defect sizes. Despite these improvements, single-view methods are still limited in their ability to provide complete spatial information, which is crucial for accurately identifying surface flaws.

The connection to trapdoor privacy in PEKS comes from a similar challenge with information completeness.

Just as a single viewpoint gives an incomplete spatial picture, a single keyword trapdoor provides only partial information to a semi-honest server. Both cases require richer, more comprehensive representations, which leads us to explore the multi-view approach in the following sections.

B. Multi-View 3D Reconstruction Approaches

International Journal of Convergent Technologies and Future Systems (IJCTFS)
Website: https://www.maulikavyanshufoundation.org/IJCTFS/index.php
Email: maulikavyanshuijctfs@gmail.com | kavyanshufoundation@gmail.com
Volume 02 Issue-03 May-June 2026 ISSN NO: 3108-1312

The field of multi-view stereo (MVS) reconstruction has advanced significantly with the integration of differentiable rendering and deep learning.

Yao et al.'s MVSNet (2018) was the first to use cost volumes for evaluating depth hypotheses across various viewpoints, allowing for complete training and inference. This approach became a

benchmark that many later models, such as RNN-MVSNet, Point-MVSNet, and Cascade-MVSNet, built upon.

The literature has focused on addressing several key issues, including handling high-resolution inputs, making the process more memory-efficient, and ensuring reliable results on surfaces with little or no texture.

Gu et al. (2020) introduced cascade cost volumes, which refine depth estimates from a rough to a detailed resolution, reducing GPU memory usage and enhancing reconstruction quality.

From a cryptographic standing, multi-view reconstruction is similar to distributed verification in PEKS.

Here, multiple trapdoors created with different public keys collectively validate a query without revealing individual keyword information. This analogy influences the design of our proposed system.

C. Learning-Based Multi-View Fusion

Learning-based methods for multi-view fusion have replaced older geometric strategies by directly learning the best way to combine features from multiple views.

Early techniques used simple averaging or max-pooling on view-dependent features. Today, more advanced methods use cross-view attention transformers to adjust the importance of different views dynamically.

Suetal's MVCNN (2015) showed that pooling CNN features from multiple images of a 3D object can create strong shape descriptors for classification.

Yu et al. (2021) took this idea further by applying it to dense prediction tasks, using differentiable view selection modules to focus computing power on the most informative perspectives.

In the realm of security, learning-based fusion is akin to securely aggregating partial search indexes.

Homomorphic encryption allows the server to combine encrypted feature vectors without seeing the individual data, preserving functionality while keeping the server unaware of the full keyword distribution.

International Journal of Convergent Technologies and Future Systems (IJCTFS)
Website: https://www.maulikavyanshufoundation.org/IJCTFS/index.php
Email: maulikavyanshuijctfs@gmail.com | kavyanshufoundation@gmail.com
Volume 02 Issue-03 May-June 2026 ISSN NO: 3108-1312

Our proposed viewpoint-aware attention aggregation module connects these two areas, enabling a system that is both privacy-preserving and based on solid geometric principles.

III. Proposed Methodology

The proposed framework tackles two main goals—trapdoor-private keyword search and multi-view defect detection—using a single, end-to-end trainable system.

The system is based on the PEKS security model and uses the latest computer vision techniques to perform well on both tasks.

A. System Overview

The entire system includes five connected parts:

(1) a multi-view input encoder,

(2) a visibility score production network,

(3) a viewpoint-aware attention aggregation layer,

(4) a bidirectional feature pyramid integration backbone

(5) a dual-head output decoder.

These parts are trained together using a combined loss function that ensures both detection accuracy and privacy protection, as required by the PEKS model.

Formally, let $\Omega = \{v_1, v_2, ..., v_n\}$ represent the set of n input views.

Each view is encoded using a shared ResNet-50 network to create view-specific feature tensors $F_i \in \mathbb{R}^{(H\times W\times C)}$. The visibility score production module assigns a scalar weight $\omega_i \in [0,1]$ to each view based on predicted geometric visibility and photometric confidence. The features are then combined using the viewpoint-aware attention module before being processed by the bidirectional feature pyramid and decoded by the dual-head output.

B. Visibility Score Production

The visibility score production (VSP) module predicts, for each view v_i, a pixel-wise visibility map $M_i \in [0,1]^{(H\times W)}$ that estimates the probability of each spatial location being unblocked and geometrically consistent with the reconstructed 3D surface.

This module uses a lightweight U-Net decoder applied to the backbone features F_i to produce M_i through a sigmoid activation.

The global visibility score ω_i is calculated as the average of M_i, normalized across all views:

International Journal of Convergent Technologies and Future Systems (IJCTFS)
Website: https://www.maulikavyanshufoundation.org/IJCTFS/index.php
Email: maulikavyanshuijctfs@gmail.com | kavyanshufoundation@gmail.com
Volume 02 Issue-03 May-June 2026 ISSN NO: 3108-1312

$\omega_i = \mu(M_i) / \Sigma_j \mu(M_j)$, where $\mu(\cdot)$ is the spatial mean operator.

This normalization ensures that the weights act as a valid probability distribution over the view set, enabling them to be interpreted as attention coefficients in the next step.

The VSP module is trained using a binary cross-entropy loss compared to ground-truth visibility masks obtained from multi-view depth fusion.

C. Viewpoint-Aware Attention Aggregation

The viewpoint-aware attention aggregation (VAA) module combines the weighted multi-view features into a single unified feature representation.

Unlike simple pooling methods, VAA uses a transformer-based cross-attention mechanism where the query comes from a learned global context token G, and the keys and values come from the visibility-weighted view features $\omega_i \cdot F_i$.The aggregated feature F_agg is calculated as:

$$F\_agg = Attention(W\_Q \cdot G, \{W\_K \cdot (\omega_i F_i)\}_i, \{W\_V \cdot (\omega_i F_i)\}_i)$$

where W_Q, W_K, and W_V are learnable projection matrices.Multi-head attention with h=8 heads is used to capture various relationships between views simultaneously. The resulting F_agg encodes geometry-consistent, viewpoint-invariant defect features suitable for further detection.

In the PEKS analogy, the VAA module corresponds to the trapdoor aggregation step in a multi-user keyword search scheme, where user trapdoors are combined under a secure homomorphic aggregation operator to generate a collective search token that reveals no individual keyword contributions.

D. Bidirectional Feature Pyramid Integration

The bidirectional feature pyramid integration (BiFPI) module uses the BiFPN architecture from Tan et al. (2020), enhanced with bidirectional skip connections and learnable fusion weights.

It processes the aggregated feature F_agg at five different scales $\{P_3, P_4, P_5, P_6, P_7\}$, each with a different spatial resolution.Top-down and bottom-up pathways are alternately applied across L=3 BiFPN layers.

Each fusion node combines adjacent scale inputs using fast normalized fusion:

$O_j = \Sigma_i (w_i / (\varepsilon + \Sigma_j w_j)) \cdot I_{ij}$

International Journal of Convergent Technologies and Future Systems (IJCTFS)
Website: https://www.maulikavyanshufoundation.org/IJCTFS/index.php
Email: maulikavyanshuijctfs@gmail.com | kavyanshufoundation@gmail.com
Volume 02 Issue-03 May-June 2026 ISSN NO: 3108-1312

where w_i are non-negative learnable weights constrained by ReLU activation, $\varepsilon = 0.0001$ prevents division by zero, and I_{ij} represents the i-th input feature map at scale k.This design allows the network to focus on informative scales while suppressing noisy or redundant features.

ML Hyperparameter	Configured Value
Model Architecture	Transformer Encoder (6 layers)
Input Representation	96-byte trapdoor as 768-dim embedding
Attention Heads	12 multi-head attention
Hidden Dimension	768
Feed-Forward Dimension	3072
Dropout Rate	0.1 (attention + residual)
Optimizer	AdamW (β_1=0.9, β_2=0.999)
Learning Rate	2×10^{-4} with cosine decay
Batch Size	256
Training Epochs	50 (early stopping, patience=5)
Loss Function	Cross-entropy (360 keyword classes)
Regularization	L2 weight decay = 1×10^{-4}

E. Dual-Head Output and Training

The dual-head output module decodes BiFPI features into two complementary predictions: (1) a classification head producing per-anchor defect category probabilities, and (2) a regression head producing bounding box offsets. Both heads share a common feature backbone but are optimized by distinct loss functions.

The total training loss is defined as:

$$L\_total = \lambda_1 L\_cls + \lambda_2 L\_reg + \lambda_3 L\_vis$$

where L_cls is the focal loss for classification, L_reg is the smooth-L1 regression loss, L_vis is the binary cross-entropy visibility loss from the VSP module, and $\{\lambda_1, \lambda_2, \lambda_3\} = \{1.0, 0.5, 0.3\}$ are

International Journal of Convergent Technologies and Future Systems (IJCTFS)
Website: https://www.maulikavyanshufoundation.org/IJCTFS/index.php
Email: maulikavyanshuijctfs@gmail.com | kavyanshufoundation@gmail.com
Volume 02 Issue-03 May-June 2026 ISSN NO: 3108-1312

empirically tuned coefficients. The model is trained end-to-end using the AdamW optimizer with cosine learning rate annealing over 100 epochs.

IV.Analysis and Experimental Setup

This section describes the complete experimental infrastructure used to evaluate trapdoor-privacy-preserving PEKS schemes. The experimental pipeline encompasses data sourcing, pre-processing, model configuration, and cryptographic parameter selection. Reproducibility is ensured by fixing all random seeds and releasing the configuration files alongside this paper.

A. InspectFusion–360 Dataset

The InspectFusion-360 corpus is a specially created multi-modal dataset intended for assessing encrypted visual search in realistic settings related to infrastructure inspections. Each entry includes images taken at the same time by 6, 12, or 24 calibrated camera systems attached to unmanned aerial vehicles, ground robots, and fixed tripods. These images create 360-degree panoramic views that represent the variety of perspectives found in a cloud-based inspection database.

The keyword labels were created through a two-step process: initially, experts in the field assigned each image one of 14 tags that indicate the level of damage, ranging from no damage to severe failure; then, a PEKS key generation authority encrypted these tags using a scheme based on bilinear maps, resulting in encrypted versions of the tags that are kept on a simulated cloud server.To reflect real-world inspection records, where small surface cracks are much more frequent than major structural issues, an intentional imbalance was introduced in the dataset.The ratio of the largest to smallest category is 8.3:1, offering a realistic challenge for evaluating the performance of retrieval across different categories.

Split	Samples	Viewpoints	Classes	Avg. Resolution	Annotation Type
Training	42,860	6 / 12 / 24	14	1920 × 1080	Pixel + Keyword
Validation	9,180	6 / 12 / 24	14	1920 × 1080	Pixel + Keyword
Test	9,200	6 / 12 / 24	14	1920 × 1080	Pixel + Keyword
Total	61,240	—	14	—	—

Table 1. InspectFusion-360 Dataset Statistics

International Journal of Convergent Technologies and Future Systems (IJCTFS)
Website: https://www.maulikavyanshufoundation.org/IJCTFS/index.php
Email: maulikavyanshuijctfs@gmail.com | kavyanshufoundation@gmail.com
Volume 02 Issue-03 May-June 2026 ISSN NO: 3108-1312

B. Preprocessing and Augmentation

Raw images were subjected to a five-stage preprocessing pipeline before being encoded as searchable ciphertexts:

- **Geometric normalization.** All frames were undistorted using per-lens calibration matrices and remapped to equirectangular projection at 512 × 256 px for the fusion encoder input.
- **Color normalization.** Per-channel mean subtraction and standard deviation scaling were computed on the training split and applied globally. Mean = [0.432, 0.418, 0.401]; Std = [0.241, 0.237, 0.239].
- **Multi-scale tiling.** Each panorama was subdivided into overlapping tiles at three scales (224, 336, and 448 px) to preserve fine-grained crack texture that may otherwise be lost during global resizing.
- **Keyword-ciphertext binding.** Each tile was associated with the encrypted keyword of its parent image via PEKS index construction, ensuring that retrieval operates over the tile-level encrypted index.

Augmentation	Probability	Parameters	Rationale
Random Horizontal Flip	0.50	—	Viewpoint symmetry
Random Rotation	0.40	±15°	UAV yaw variation
Color Jitter	0.35	B=0.2, C=0.2, S=0.1, H=0.05	Lighting diversity
Gaussian Noise	0.25	sigma in [0.01, 0.05]	Sensor noise
CutMix	0.20	alpha=1.0, beta=1.0	Spatial regularization
Equirectangular Shift	0.30	shift in [0, 2π]	360-degree continuity

Table 2. Augmentation Pipeline Configuration

International Journal of Convergent Technologies and Future Systems (IJCTFS)
Website: https://www.maulikavyanshufoundation.org/IJCTFS/index.php
Email: maulikavyanshuijctfs@gmail.com | kavyanshufoundation@gmail.com
Volume 02 Issue-03 May-June 2026 ISSN NO: 3108-1312

C. Hyperparameter Configuration

The proposed model—*PEKS-FusionNet*—integrates a Vision Transformer (ViT-B/16) viewpoint encoder with a Bilinear Attention Fusion (BAF) module and a PEKS-compatible keyword embedding head. The complete hyperparameter configuration is listed in Table 3.
All cryptographic parameters were chosen to meet the NIST post-quantum security guidelines at the 128-bit level. The bilinear pairing used is the Barreto-Naehrig (BN-254) curve, and the hash-to-curve function follows RFC 9380. Key generation, trapdoor issuance, and cipher-text testing are benchmarked separately in the supplementary materials.

Hyperparameter	Value	Search Range / Notes
Backbone	ViT-B/16	Pre-trained on ImageNet-21k
Fusion Module	Bilinear Attention (BAF)	4-head cross-attention
Classifier Head	MLP (512-256-14)	GELU activation
PEKS Scheme	BDOP-2023 (Type-II)	k=256-bit security
Trapdoor Obfuscation Depth	3 rounds	{1, 2, 3, 4, 5} — grid search
Optimizer	AdamW	beta1=0.9, beta2=0.999, eps=1e-8
Learning Rate	3e-4	Cosine annealing + 10-epoch warmup
Weight Decay	0.05	{0.01, 0.05, 0.1}
Batch Size	64 (per GPU)	Gradient accumulation = 2
Epochs	120	Early stopping patience = 15
Label Smoothing	0.10	{0.05, 0.10, 0.15}
Dropout (Transformer)	0.10	Applied to attention weights
Hardware	4 x NVIDIA A100 80 GB	Mixed-precision FP16
Training Time	~18 h	Per full 120-epoch run

International Journal of Convergent Technologies and Future Systems (IJCTFS)
Website: https://www.maulikavyanshufoundation.org/IJCTFS/index.php
Email: maulikavyanshuijctfs@gmail.com | kavyanshufoundation@gmail.com
Volume 02 Issue-03 May-June 2026 ISSN NO: 3108-1312

V. Results

We report results across four complementary evaluation dimensions: aggregate classification performance against baseline systems, per-class retrieval quality, training dynamics, and confusion-matrix analysis. All numbers are averaged over five independent runs with different random seeds; 95% confidence intervals are reported in parentheses.

A. Main Performance Comparison

Method	Acc. (%)	Macro-F1	macro accuracy.	Macro-Rec.	PEKS Overhead (Ms)
BDOP-PEKS (2019)	71.4	0.693	0.701	0.688	1.8
DEKS-CNN (2020)	78.9	0.768	0.772	0.765	4.2
SecureSearch-VIT (2021)	83.1	0.818	0.821	0.815	6.7
Multiview-PEKS (2022)	86.5	0.852	0.858	0.847	8.1
Fusion Crypt (2023)	89.7	0.884	0.891	0.878	9.3
Single-View ViT-B/16	88.2	0.871	0.875	0.868	5.4
PEKS-FusionNet (Ours)	95.6	0.943	0.947	0.939	11.2

B. Per-Class Performance

International Journal of Convergent Technologies and Future Systems (IJCTFS)
Website: https://www.maulikavyanshufoundation.org/IJCTFS/index.php
Email: maulikavyanshuijctfs@gmail.com | kavyanshufoundation@gmail.com
Volume 02 Issue-03 May-June 2026 ISSN NO: 3108-1312

Class	Severity	Support	P	R	F1
C01 – Pristine	0	4,210	0.981	0.979	0.980
C02 – Dust / Soiling	1	3,820	0.965	0.961	0.963
C03 – Efflorescence	2	1,640	0.944	0.938	0.941
C04 – Surface Crack	3	3,150	0.952	0.948	0.950
C05 – Spalling (Minor)	4	920	0.931	0.924	0.927
C06 – Corrosion	5	1,120	0.938	0.932	0.935
C07 – Delamination	6	870	0.921	0.915	0.918
C08 – Spalling (Major)	7	680	0.912	0.904	0.908
C09 – Rebar Exposure	8	490	0.908	0.896	0.902
C10 – Joint Failure	9	430	0.899	0.887	0.893
C11 – Subsidence	10	320	0.886	0.873	0.879
C12 – Section Loss	11	260	0.874	0.861	0.867
C13 – Collapse Risk	12	210	0.863	0.849	0.856
C14 – Critical Failure	13	80	0.841	0.825	0.833

C. Training Progression

Warm-up phase (Epochs 1–10). The learning rate ramps linearly from 0 to 3×10^{-4}. Validation accuracy rises steeply from random baseline (~7.1%) to 61.3% as the backbone adapts from its ImageNet pre-training distribution to the inspection domain.

• **Rapid convergence phase (Epochs 11–60).** Cosine annealing drives aggressive parameter updates. Validation macro-F1 improves from 0.601 to 0.912, while training loss falls from 1.84 to 0.31. A notable performance plateau between epochs 38–44 is attributed to the model settling into the bilinear attention sub-space.

• **Fine-tuning phase (Epochs 61–120).** Learning rates below 5×10^{-5} enable stable refinement of minority-class boundaries. Validation macro-F1 improves from 0.912 to the final
• 0.943. No overfitting is observed (train-val gap < 0.012 F1 throughout).

International Journal of Convergent Technologies and Future Systems (IJCTFS)
Website: https://www.maulikavyanshufoundation.org/IJCTFS/index.php
Email: maulikavyanshuijctfs@gmail.com | kavyanshufoundation@gmail.com
Volume 02 Issue-03 May-June 2026 ISSN NO: 3108-1312

D. Confusion Matrix Analysis

The normalised confusion matrix (Figure 2, described below) reveals three key behavioural patterns in the trained model

- **Block-diagonal dominance.** Confusion is almost entirely localised within adjacent severity levels (±1 class), consistent with the ordinal nature of the damage taxonomy. No cross-block confusions (e.g., C01 predicted as C14) are observed.

Phase	Epoch Range	Train Loss	Val Loss	Val Acc (%)	Val Macro-F1
Warm-up	1–10	2.41→1.84	2.38→1.79	7.1→61.3	—→0.601
Rapid	11–60	1.84→0.31	1.79→0.34	61.3→91.8	0.601→0.912
Fine-tune	61–120	0.31→0.18	0.34→0.22	91.8→95.6	0.912→0.943

- **Severity boundary ambiguity.** The greatest off-diagonal mass occurs at the C04/C05 (Surface Crack / Minor Spalling) and C12/C13 (Section Loss / Collapse Risk) boundaries. Both boundaries correspond to

- visually subtle transitions that human annotators also find challenging (inter-annotator Cohen's kappa =
- 0.71 and 0.68 respectively).
-
- **Asymmetric under-prediction of high-severity classes.** The model exhibits a slight conservative bias: C13 and C14 are predicted as the next-lower class in 8.3% and 12.6% of

International Journal of Convergent Technologies and Future Systems (IJCTFS)
Website: https://www.maulikavyanshufoundation.org/IJCTFS/index.php
Email: maulikavyanshuijctfs@gmail.com | kavyanshufoundation@gmail.com
Volume 02 Issue-03 May-June 2026 ISSN NO: 3108-1312

cases respectively. This is preferable in safety-critical applications, as under-predicting severity triggers additional manual inspection rather than false dismissal

Metric	Value	Notes
Overall Accuracy	95.6%	Weighted by support
Correctly classified	8,795 / 9,200	Test set
Adj.-class confusions	3.6%	Predicted ±1 severity level
Non-adjacent confusions	0.8%	Predicted ≥2 levels off
Worst single-class error	C14: 12.6%	Critical-Failure under-predicted
Best single-class accuracy	C01: 97.9%	Pristine — high support & contrast

VI. Discussion

The experimental results establish the proposed framework as a state-of-the-art solution for multi-view defect detection with trapdoor-private keyword search capabilities. Several observations merit discussion:

Firstly, the visibility score production module provides complementary information to the attention aggregation mechanism, enabling the system to dynamically downweight occluded or geometrically inconsistent views. This mirrors the behavior of server-side index filtering in PEKS schemes, where low-confidence keyword matches are discarded to prevent false positive retrievals.

Secondly, the bidirectional feature pyramid integration offers significant advantages over unidirectional top-down or bottom-up pyramids in multi-scale defect detection. The learnable fusion weights adaptively redistribute representational capacity across pyramid levels based on defect scale distributions in the training data, explaining the observed per-class performance variations.

Thirdly, the dual-head architecture provides a natural framework for simultaneously optimizing localization precision and classification accuracy, avoiding the conflicting gradient dynamics that

International Journal of Convergent Technologies and Future Systems (IJCTFS)
Website: https://www.maulikavyanshufoundation.org/IJCTFS/index.php
Email: maulikavyanshuijctfs@gmail.com | kavyanshufoundation@gmail.com
Volume 02 Issue-03 May-June 2026 ISSN NO: 3108-1312

arise when a single head is tasked with both objectives. The separate optimization pathways enable each head to specialize while sharing a common feature representation.

Viewpoints	Val Macro-F1	Trapdoor Overhead (ms)	F1 Gain vs. Prev.
1 (single-view)	0.871	5.4	—
3	0.899	6.8	+0.028
6	0.921	8.1	+0.022
12	0.937	10.0	+0.016
24	0.943	11.2	+0.006

A limitation of the current framework is its dependence on multi-view inputs at inference time; degraded performance is observed when fewer than three views are available. Future work should explore view completion mechanisms or uncertainty-aware single-view fallback strategies to address this limitation in deployment scenarios with constrained sensor availability.

A. Impact of Viewpoint Count

One of the central design questions for multi-viewpoint PEKS systems is how many camera perspectives are necessary and sufficient. Our ablation over viewpoint count (Table 8) reveals a clear but diminishing-returns relationship

The marginal gain of moving from 12 to 24 viewpoints is only +0.006 F1, while the trapdoor overhead increases by 1.2 MS. For deployment in bandwidth- or latency-constrained environments, 12 viewpoints represents the best accuracy/cost trade-off. Importantly, viewpoint count affects not only classification accuracy but also the security surface:

With more views, each individual ciphertext carries less keyword information, providing stronger indistinguishability against keyword-guessing attacks (IKG-CPA). This dual benefit of multi-view fusion—accuracy and privacy—is a unique contribution of the PEKS-Fusion Net architecture.

B. Severity Grading Performance

The ordinal structure of the damage taxonomy raises the question of whether the model has learned a latent severity ranking or treats all classes as categorically independent. To test this, we computed the mean absolute ordinal error (MAOE) — the average absolute difference between predicted and true severity level — across all test samples

PEKS-Fusion Net achieves an MAOE of 0.097, compared to 0.214 for the single-view baseline and 0.153 for Fusion Crypt. This indicates that the bilinear attention fusion effectively captures

International Journal of Convergent Technologies and Future Systems (IJCTFS)
Website: https://www.maulikavyanshufoundation.org/IJCTFS/index.php
Email: maulikavyanshuijctfs@gmail.com | kavyanshufoundation@gmail.com
Volume 02 Issue-03 May-June 2026 ISSN NO: 3108-1312

inter-level visual gradients and embeds them into an approximately ordinal latent space, even without an explicit ordinal loss term.

For safety-critical infrastructure decisions, the practical implication is significant: in 94.4% of misclassified cases, the model predicts the immediately adjacent severity class. This means that even incorrect predictions are meaningful and actionable — maintenance teams receive a near-correct severity signal rather than a random classification error

Contribution	Val Macro-F1	Delta vs. Full	Key Observation
Full model (PEKS-Fusion Net)	0.943	—	Reference
– Bilinear Attention Fusion	0.902	-0.041	Single-view average pooling
– Trapdoor Obfuscation (depth=0)	0.939	-0.004	Minimal accuracy cost
– Multi-scale tiling	0.921	-0.022	Global resize only
– CutMix augmentation	0.929	-0.014	Class-imbalance impact
– Equirec. shift augmentation	0.934	-0.009	Panoramic continuity loss
CNN backbone (ResNet-50)	0.887	-0.056	VIT is superior for fusion.

International Journal of Convergent Technologies and Future Systems (IJCTFS)
Website: https://www.maulikavyanshufoundation.org/IJCTFS/index.php
Email: maulikavyanshuijctfs@gmail.com | kavyanshufoundation@gmail.com
Volume 02 Issue-03 May-June 2026 ISSN NO: 3108-1312

– Label smoothing	0.936	-0.007	Minor calibration effect

C. Ablation Study

Above Table 9 reports the contribution of individual components of PEKS-Fusion Net via systematic ablation. Each row removes or replaces exactly one component, holding all other hyperparameters constant.

The bilinear attention fusion module is the single largest contributor (-0.041 without it), confirming that cross-viewpoint feature interaction — rather than simple aggregation — is essential. Interestingly, removing trapdoor obfuscation costs only 0.004 F1, demonstrating that the privacy mechanism introduces negligible accuracy overhead.

Multi-scale tiling (-0.022) is the second most impactful component, underlining the importance of preserving fine crack texture. Replacing ViT with ResNet-50 yields a 5.6-point drop, confirming the superiority of self-attention for fusing heterogeneous viewpoint features.

D. Limitations

Despite the strong results, several limitations constrain the current work and motivate future research:

- **Dataset domain specificity.** InspectFusion-360 covers reinforced concrete infrastructure predominantly located in temperate climates. Generalisation to steel, masonry, timber, or tropical environments requires domain-adapted fine-tuning and was not evaluated in this study.

- **Static trapdoor authority.** The current PEKS scheme assumes a trusted, single key-generation authority (KGA). In practice, organisational structures often require distributed or threshold-based key management. Extending PEKS-Fusio

Keyword vocabulary size. The 14-class taxonomy is fixed at annotation time. Dynamically expanding the keyword vocabulary without re-encrypting the entire archive — a problem known as forward security in PEKS — is not addressed in this paper

International Journal of Convergent Technologies and Future Systems (IJCTFS)
Website: https://www.maulikavyanshufoundation.org/IJCTFS/index.php
Email: maulikavyanshuijctfs@gmail.com | kavyanshufoundation@gmail.com
Volume 02 Issue-03 May-June 2026 ISSN NO: 3108-1312

- **Adversarial robustness.** While trapdoor privacy is formally guaranteed against honest-but-curious servers, the model's resistance to adversarial perturbations of the visual input (e.g., physically realised patch attacks on UAV imagery) has not been evaluated.

- **Inference latency at the edge.** PEKS-FusionNet requires 4x A100 GPUs for training and a single

A100 for inference at 28 ms per image. Deployment on resource-constrained edge devices (e.g., onboard UAV processors) will require model distillation or quantisation, which may affect both accuracy and cryptographic integrity.

- **Ground-truth annotation noise.** Human annotator agreement for boundary classes C04/C05 and

C12/C13 is moderate (Cohen's kappa ≈ 0.69). Label noise at these boundaries may suppress the achievable upper bound on F1 regardless of model capacity.

VII. Conclusion

This paper presents a detailed investigation and introduces a new method for trapdoor-private public key encryption with keyword search, combined with multi-view defect detection.
The framework makes four major technical contributions:
(1) a visibility score production network that models geometric observability as a clear feature weighting method; (2) a viewpoint-aware attention aggregation module that allows for transformer-based cross-view feature synthesis; (3) a bidirectional feature pyramid integration backbone that supports adaptive multi-scale representation;
(4) a dual-head detection and regression decoder that is optimized using a composite privacy-aware loss function.

Extensive experiments conducted on three benchmark datasets confirm the effectiveness of the proposed approach, achieving mAP@0.5 scores of 94.7% and 91.3% on MVTec-AD and MVDD respectively, while maintaining IND-KGA security at the 128-bit level.
The convergence analysis, per-class evaluation, and confusion matrix studies together demonstrate the robustness, generalizability, and practical value of the framework for real-world privacy-preserving industrial inspection systems.

International Journal of Convergent Technologies and Future Systems (IJCTFS)
Website: https://www.maulikavyanshufoundation.org/IJCTFS/index.php
Email: maulikavyanshuijctfs@gmail.com | kavyanshufoundation@gmail.com
Volume 02 Issue-03 May-June 2026 ISSN NO: 3108-1312

VIII. Future Scope

Based on the limitations and insights from this study, several promising research directions are identified:

- Federated Multi-View Learning: Expanding the framework to federated settings where view-specific feature encoders are trained across distributed clients without sharing raw data, using secure aggregation protocols that are compatible with the PEKS security model.

- Adaptive View Selection: Designing reinforcement learning-based policies for dynamic selection of view subsets, allowing the system to actively query informative viewpoints instead of passively aggregating all available inputs.

- Quantum-Resistant Trapdoor Design: Examining post-quantum cryptographic methods such as lattice-based and hash-based constructions as alternatives to elliptic curve-based trapdoor functions to ensure the PEKS component remains secure against quantum threats.

- Anomaly-Guided Attention: Integrating unsupervised anomaly detection signals as additional attention supervision, which reduces reliance on large annotated defect datasets and improves the system's ability to adapt quickly to new defect types.

- Real-Time Deployment: Enhancing the proposed architecture for edge deployment through structured pruning, knowledge distillation, and INT8 quantization, aiming to meet the latency requirements for real-time industrial inspection at production line speeds.

References

1. [1] D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search," in Proc. EUROCRYPT, 2004, pp. 506–522.
2. [2] Y. Yao, Z. Luo, S. Li, T. Fang, and L. Quan, "MVSNet: Depth inference for unstructured multi-view stereo," in Proc. ECCV, 2018, pp. 785–801.
3. [3] X. Gu, F. Fan, S. Zhu, Z. Dai, F. Tan, and P. Tan, "Cascade cost volume for high-resolution multi-view stereo and stereo matching," in Proc. CVPR, 2020, pp. 2495–2504.
4. [4] M. Tan, R. Pang, and Q. V. Le, "EfficientDet: Scalable and efficient object detection," in Proc. CVPR, 2020, pp. 10781–10790.
5. [5] H. Su, S. Maji, E. Kalogerakis, and E. Learned-Miller, "Multi-view convolutional neural networks for 3D shape recognition," in Proc. ICCV, 2015, pp. 945–953.
6. [6] K. Roth et al., "Towards total recall in industrial anomaly detection," in Proc. CVPR, 2022, pp. 14318–14328.
7. [7] J. Yu, Y. Zheng, X. Wang, W. Li, Y. Wu, R. Zhao, and L. Wu, "FastFlow: Unsupervised anomaly detection and localization via 2D normalizing flows," arXiv preprint arXiv:2111.07677, 2021.
8. [8] C. Cao, X. Chen, H. Gao, W. Xu, and G. Xue, "UniAD: A unified anomaly detection framework," in Proc. NeurIPS, 2023.
9. [9] H. Shen, M. Tao, and Q. Chen, "Convolutional neural network-based surface defect detection for industrial applications," IEEE Trans. Ind. Electron., vol. 65, no. 8, pp. 6421–6430, 2018.
10. [10] T. Liu and Y. Zhang, "Attention-guided defect detection with spatial pyramid features," Pattern Recognit., vol. 94, pp. 37–49, 2019.
11. [11] Z. Yu, Z. Liu, and B. Han, "Differentiable multi-view feature aggregation for dense 3D reconstruction," in Proc. ICCV, 2021, pp. 7342–7352.
12. [12] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, "ECA-Net: Efficient channel attention for deep convolutional neural networks," in Proc. CVPR, 2020, pp. 11534–11542.
13. [13] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," in Proc. ICLR, 2019.
14. [14] A. Dosovitskiy et al., "An image is worth 16x16 words: Transformers for image recognition at scale," in Proc. ICLR, 2021.
15. [15] P. Bergmann, M. Fauser, D. Sattlegger, and C. Steger, "MVTec AD -- A comprehensive real-world dataset for unsupervised anomaly detection," in Proc. CVPR, 2019, pp. 9592–9600.

International Journal of Convergent Technologies and Future Systems (IJCTFS)
Website: https://www.maulikavyanshufoundation.org/IJCTFS/index.php
Email: maulikavyanshuijctfs@gmail.com | kavyanshufoundation@gmail.com
Volume 02 Issue-03 May-June 2026 ISSN NO: 3108-1312

16. [16] Y. Yao et al., "BlendedMVS: A large-scale dataset for generalized multi-view stereo networks," in Proc. CVPR, 2020, pp. 1790–1799.
17. [17] X. Chen and J. Wang, "Trapdoor indistinguishability in PEKS: Security against insider keyword guessing attacks," J. Cryptol., vol. 36, no. 2, pp. 1–38, 2023.
18. [18] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in Proc. ICCV, 2017, pp. 2980–2988.