Perceptual Colour-based Geolocation of Human Trafficking Images for Digital Forensic Investigation

Jessica Herrmann*, Opeyemi Bamigbade†, John Sheppard†, Mark Scanlon*
*Forensics and Security Research Group, School of Computer Science, Univeristy College Dublin, Ireland.
herrmann.jessica@protonmail.com, mark.scanlon@ucd.ie
†Department of Computing and Mathematics, South East Technological University, Waterford, Ireland.
opeyemi.bamigbade@postgrad.wit.ie, john.sheppard@setu.ie

***Abstract*—This paper investigates the effectiveness of colour-based descriptors in Content-Based Image Retrieval (CBIR) and examines the impact of various parameters on image matching accuracy. The aim is to improve image retrieval methods to support digital forensic investigators in human trafficking cases. Colour values are used as key components to describe specific image characteristics, and the technique is evaluated on the Hotels-50K dataset. The method achieved a Top-50 accuracy of over 95%, enabling efficient data triage and significantly reducing the volume of images to be examined. Using 2 colour descriptors is found to optimise the balance between information richness and dimensionality reduction. Performance is further improved by optimised image selection, reducing false-positive rates, and increasing robustness. The approach demonstrates potential in advancing image analysis tools in human trafficking investigations and other contexts, opening new avenues for using colour values in crime detection and image data analysis. Future research may refine the Euclidean distance method used in the image similarities measure by introducing weighted distance measurements to reduce the impact of common colour values, and investigate lighting and saturation effects.**

***Index Terms*—Content-Based Image Retrieval (CBIR), Colour Descriptors, Computer Vision, Human Trafficking, Artificial Intelligence, Multimedia Geolocation**

I. Introduction

Human trafficking, a complex and serious crime, affects millions of people worldwide, transcending age, gender, and background, and causes severe personal, communal, and societal harm. It involves the illegal trade of individuals through deception, violence, extortion, or exploitation, leading to various forms of abuse, including forced labour, sexual exploitation, and organ trafficking. The European Union employs “Directive 2011/36/EU” to combat this crime, setting rules for offences and sanctions, prioritising prevention, and victim protection with a gender perspective [1].

European data collection on trafficking progress occurs biannually, revealing deficits in data and technology-assisted investigative methods. Identification of trafficking victims typically relies on location analysis, phone numbers, and textual content, while image analysis, especially involving image content, is underused or limited to exact matches. Stylianou et al. [2] underscored the importance of images in fighting human trafficking and the potential of Content-Based Image Retrieval (CBIR). CBIR uses visual characteristics to analyse and find similarities in images, employing advanced technologies such as neural networks. Despite practical challenges in data collection and processing, the creation of specialised datasets, e.g., Hotels-50K [3] and Hotel-ID [4], has driven significant research progress in image matching for potential crime locations. Similar hotel images are often posted by Europol’s “Stop Child Abuse – Trace an Object” project to crowdsource the location[1]. In many cases, the section of the images that contain the victim is hidden to protect privacy and comply with legal regulations.

The increasing volume of data underscores the need for AI-based solutions, particularly for digital forensic laboratories that handle large datasets [5]. However, artificial intelligence (AI) technology remains in its early stages, with significant gaps between research on analysis methods and practical applications for law enforcement [6]. These tools must not only produce valid results, but also comply with legal, ethical, and practical standards [7].

The ethical considerations surrounding AI have been formally recognised at the EU level with the enactment of the AI Act, which came into force in August 2024 [8]. This comprehensive legislation aims to ensure that the benefits of AI are realized while adhering strictly to ethical principles and safety standards. The AI Act establishes a unified regulatory framework that promotes the development of secure and trustworthy AI systems and regulates their responsible use across the private and public sectors.

There is a growing trend within the digital forensic community toward approaches aided by AI [6, 9], indicating that intelligently automated digital evidence processing is gaining in popularity as a technique [10]. In an environment of increasing data volumes, CBIR shows potential for improving image analysis during digital investigation. Specifically, the use of colour as a key feature in CBIR [11]. By incorporating CBIR into investigations, law enforcement can better analyse photos and videos, revealing patterns linked to human trafficking that might be missed by traditional methods.

This paper contributes to CBIR research by deeply exploring the role of colour in content-based image matching. Through empirical analysis using the Hotels-50K dataset, the study highlights the effectiveness of colour features in distinguishing and retrieving relevant images. These findings have practical implications for visual content analysis, particularly in human

[1]https://www.europol.europa.eu/stopchildabuse

trafficking investigations, offering insights for the design and implementation of CBIR systems that leverage colour information.

A. Research Questions

This research evaluates a colour matching procedure for hotel image identification within the Hotels-50K dataset, focussing on its accuracy in assigning images to the correct hotels based on extracted colour values. The study examines the optimal number of colour values for maximising matching accuracy and precision, with the aim of assessing the procedure's performance and providing insights for improving image-based hotel identification.

The research focusses on the following questions:

- How effective and accurate can a matching method based on colour values identify hotel rooms in a large dataset of images from different sources?
- What influence does the number of extracted colour values have on the accuracy and precision of the colour descriptors in the matching procedure?
- What influence can colour values as hard descriptors have on reducing the search field in the identification of hotel rooms?

A hierarchical analysis of the Top-K accuracy at a 95% threshold is intended to achieve data triage, effectively narrowing the search field and enabling a more targeted image analysis. In this context, 'hard descriptors' refer to specific and well-defined features that provide consistent information about an image's color composition, allowing for precise matching and identification.

II. Literature Review

A. Hotel Image Datasets

Indoor scene recognition presents unique challenges, and existing methods developed for outdoor scenes, such as landmark recognition, are not well suited to distinguish between similar images [3]. In 2015, Stylianou et al. [2] addressed these specific challenges with the main objective of supporting investigations and prosecutions in the context of criminal activities related to sex trafficking. Here, identification of hotel rooms served as a key link to the locations of the victims [2]. For this purpose, the researchers created a database of hotel room images used for marketing and travel sales, e.g., via the Expedia Affiliate Network API, and developed a smartphone crowd-sourcing app called "TraffickCam", through which travellers could upload their own images of hotel rooms [2].

From the combined approaches, the Hotels-50K dataset was released in 2019, which consists of a total of 1,027,871 images from 50,000 unique hotels around the world [3]. However, travel website images, which are often professionally taken under ideal conditions, are often visually very different from investigation images. In contrast, crowd-sourced images, although less frequent, have more similarities to images from genuine queries [3].

B. Related Work

In 2015, Stylianou et al. [2] compared SIFT matching with Convolutional Neural Networks (CNNs) for hotel room image matching. SIFT performed well in exact matches, but faced limitations due to a small dataset and lack of real-world variation. In 2019, Stylianou et al. [12] improved results with the ResNet-50 model, enhancing hotel room image recognition.

In 2020, Xuan et al. [13] introduced "easy positive" mining to generate embeddings, aligning similar images and distancing dissimilar ones. A 2021 study expanded this by optimising hard negative triplets in deep metric learning, acknowledging their importance where distance metrics fail to capture semantic similarity. Xuan et al. [14] introduced a novel loss function and a "triplet diagram" to improve optimisation.

Subsequent research focused on hotel recognition, developing a loss function adapted for better hotel representation [15]. In 2021, Kamath et al. [4] combined Convolutional Neural Networks (CNNs) with Graph Neural Networks (GNNs) for a more comprehensive image analysis, while a study in 2022 by Black et al. [16] focused on interpretability, introducing visualisation methods to identify image similarity in police investigations.

Saikia et al. [17] highlighted the value of colour-based image search for its rich information and invariance to pixel shifts and rotations, also emphasising the importance of texture features in complex scenarios. Their algorithm combined local and global scene content to create robust descriptors for colour, texture, and perspective changes, using CNN features from various colour spaces.

Compared to existing methods that aim for exact matches, this research focusses specifically on investigating a colour-based approach that aims to identify optimal parameters for hard descriptors. The objective is to use colour descriptors to provide effective data triage, which is particularly valuable for police practice. In many forensic scenarios, effective data triage is essential, as more complex models are time- and resource-intensive. By reducing the image volume, the use of complex models can be focused specifically on the most relevant images.

C. Content-Based Image Retrieval

As image collections grow, CBIR plays a crucial role in the efficient classification and retrieval of digital images by evaluating visual attributes like colour, texture, and shape, making it ideal for searching large image databases. Its applications span fields such as medical imaging, geoinformation, and biodiversity research [18, 19, 20]. However, CBIR techniques must be tailored to each field's unique needs. In forensic applications, CBIR methods often encounter limitations due to factors such as noise, lighting variations, and fluctuating image quality, which impair their accuracy. Techniques like SIFT or CNNs, which perform well under controlled conditions, tend to lose reliability in real-world investigations and frequently result in false positives. Colour values have proven to be more robust against variations in lighting and image quality, providing a method to capture image similarities more accurately even in demanding forensic scenarios. For example, medical images are often high-resolution 3D, while surveillance images are low-resolution and noisy. Accuracy requirements also vary, with fields such as medical imaging and forensic analysis demanding high precision, particularly in legal contexts.

In law enforcement, ensuring interpretability and reliable documentation is vital [5]. Legal proceedings, especially the preservation of evidence, introduce challenges that differ between legal systems. The Artificial Intelligence Act of the EU recognises the importance of AI in law enforcement and its potential impact on fundamental rights [21]. Combining handcrafted descriptors with deep learning approaches can enhance CBIR performance, offering flexibility, lower computational needs, and improved interpretability, making them valuable supplements to CNN-extracted features.

1) CBIR Global and Local Features: CBIR feature extraction methods are generally divided into global and local approaches. Global methods extract features from the entire image, while local methods focus on smaller sections, making them more robust to geometric deformations and lighting changes.

Global methods rely on statistical properties such as average hue, colour variance, and features such as mean, standard deviation, or variance of colour values. Although efficient, this approach may miss finer image details. Histograms, commonly used to assess colour distribution, are robust to pixel shifts and rotation, focusing on relative colour distribution rather than absolute values [17]. This robustness addresses the challenges encountered in previous scene-based approaches using the Hotel-50K dataset [4]. Histograms can be created for individual colour channels or combined into a single channel.

Local methods, on the other hand, capture colour information in specific image regions. Colour textures describe characteristic features of different areas, offering greater accuracy and discrimination. Despite the success of colour histograms and shape descriptors in certain applications, they face both theoretical and practical issues, particularly the *curse of dimensionality* [22].

III. Methodology

This study follows a two-step approach to optimize and evaluate image matching parameters. First, empirical observation is conducted using a subset of the Hotels-50K dataset to identify optimal colour parameters for the matching procedure. Various parameter configurations are tested, with the selection guided by systematic analysis.

Next, the identified colour parameters are applied in a CBIR experiment. The performance of the matching procedure is evaluated by analysing the impact of the selected parameters on retrieval accuracy, as detailed in Section IV. The results are documented for each hotel in the dataset, providing insight into how the choices of colour parameters influence the quality of the image matching.

The empirical study will test the following hypothesis:

- The number of extracted colour values significantly impacts the accuracy of colour descriptors in the matching procedure, with the assumption that increasing the number of values improves matching accuracy.
- Matching performance is directly correlated with data quality, with the assumption that higher-quality data significantly enhances matching outcomes.
- Colour values as hard descriptors are hypothesised to effectively narrow the search space, thereby restricting potential matches and improving search efficiency.

The results of the matching will be evaluated and compared with the hypotheses. This will encompass the precision of the matching and any deviations from the anticipated values.

IV. Empirical Design

This analysis outlines the CBIR process and establishes the basis for the planned experiments. Insights from this chapter inform the systematic variation of different parameters in later experiments. The detailed analysis and results' presentation will serve as the groundwork for subsequent analyses and optimisations within this work.

A. Content Based Image Retrieval Process

For any problem in CBIR, the solution starts with image analysis and feature definition. This paper utilises colour features extracted from images in the dataset subset to facilitate image matching. Colour descriptors are used as the primary feature for representation. These descriptors are integral to limiting the search space and improving the precision of the matching procedure.

The performance of CBIR is evaluated using the top-k retrieval metric, where the system returns the k most similar images ranked by colour similarity. The top-k accuracy indicates the proportion of relevant matches among the first k results, serving as a key measure of the effectiveness of retrieval and the impact of the chosen colour parameters on system performance [22].

B. Dataset Configurations

The `data_source` level of the dataset contains two folders. The `travel_website` folder, which contains a collection of images from hotel websites, and the `traffickcam` folder, which contains user-contributed images from TraffickCam. Index CSV files are available for the identification of the individual images. The `chain_id` describes the hotel chain, so that the respective subfolder with the corresponding `hotel_id` can be assigned to a certain chain.

The hotels examined represent diverse hotel chains, showcasing a range of characteristics. The number of available images within the `travel_website` and `traffickcam` folders exhibited considerable variation. In `travel_website` folders, images ranged from 9 to 24, while `traffickcam` folders contained 1 to 8 images. As anticipated, `travel_website` images exhibited higher quality and more professional presentation compared to those in `traffickcam` folders, as can be seen in Figure 1. The latter often featured foreign objects and hotel guests in the pictures.

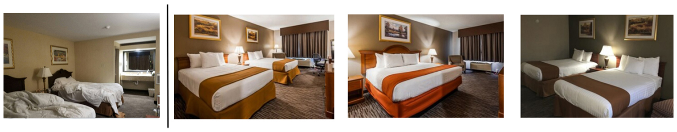

Fig. 1. On the left side: Image from the TraffickCam folder of hotel 4824, On the right side: Three images from the travel websites folder of hotel 4824

Empirical observation was carried out with a "5-hotels" subset consisting of five randomly selected hotel folders.

C. Image Preparation

Following the examination of the dataset, the initial step in CBIR involves image preparation. A potential approach is to transform colour representation into alternative colour spaces, e.g., RGB to HSV or HLS. These spaces separate the brightness and saturation from the pure colour value, allowing for colour similarity assessment based on independent colour properties, unaffected by brightness or saturation variations. An investigation into whether considering brightness or saturation as individual values may lead to insignificant differences in brightness that affect similarity ranking is deferred. The primary focus now is on the number of extracted colour values and their impact on matching precision. As a result, images are retained in their original RGB colour space.

D. Colour Palette Extraction Technique

The colour value extraction was extracted using Pylette[2], which extracts colour palettes from images using various clustering methods. In addition to RGB, HLS and HSV colour modes, Pylette supports colour quantisation using the K-Means and Median-Cut algorithms, both of which can be used to cluster similar data points. The purpose of clustering is to identify and consolidate patterns in the data. K-Means is a general clustering algorithm that is used in various fields of application to organise data into predefined clusters. Each data point is assigned to the nearest cluster centre. K-Means is widely used in data analysis, machine learning, and colour quantisation.

Hu et al. [23] compared these methods in 2013, concluding that K-Means offers better palette quality and computational efficiency, although it is more susceptible to variations due to initialisation [23]. Tests on the "5-hotels" subset confirmed this, with the Median-Cut algorithm producing consistent results, while K-Means showed slight deviations. Thus, the more robust Median-Cut was selected for further testing.

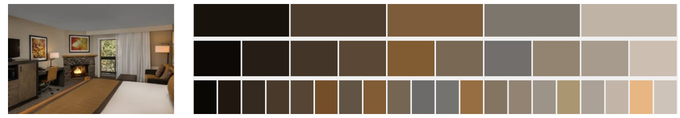

Fig. 3. Hotel 398. On the left: Image 6263111.jpg from the travel website folder, on the right: Colour palette containing 5, 10 and 20 values.

The number of colour values to be extracted can be set manually in Pylette. Initially, 5, 10 and 20 colour values were extracted per image.

Figures 2 and 3 suggest that specific colour details in palettes become distinctly visible with an extraction of at least 20 colour values or more. For example, in Figure 2 the pink of the cushion becomes prominently visible only with a higher number of extracted colour values, and the blue tones are reproduced more faithfully. In contrast, the second example image in Figure 3 exhibits barely perceptible differences compared to palettes with fewer extracted colour values, despite increased extractions. This palette primarily comprises various brown tones, which could explain the minimal visual distinctions in the created colour palettes, even with a higher number of extracted colour values.

A 2D scatter plot was used to visualise the extracted colour values, similar to histograms in representing the colour distribution but plotting individual colour values as points in a coordinate system. This allows for detecting shifts or deviations in colour values.

With Matplotlib[3], the extracted RGB values were used to create the scatter plots. NumPy[4] was used for array manipulation, and the RGB values were scaled to the range [0, 1] for consistency. The scatter plots depicted the R and G values on the corresponding axes within this range.

The scatter plots shown in Figure 4, depicting 5, 10, and 20 extracted colour values, reveal interesting patterns regarding the variety and distribution of colours. With 20 extracted values, the scatter plots exhibit a larger dispersion, indicating greater variation and colour diversity. This broadens the representation of image content. Conversely, scatter plots with only five extracted colour values display limited colour variety, especially challenging for images within similar colour ranges. These plots may not capture subtle colour differences,

[2]https://github.com/qTipTip/Pylette

[3]https://matplotlib.org/

[4]https://numpy.org/

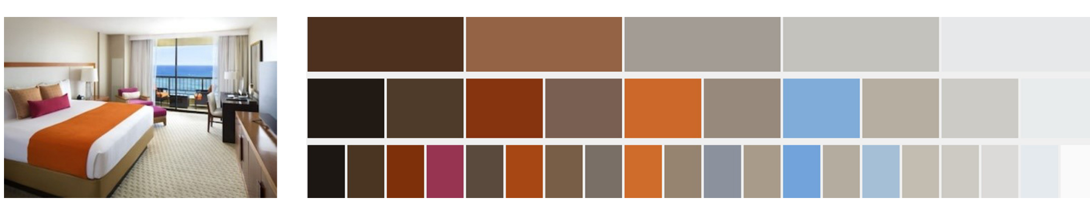

Fig. 2. Hotel 451. On the left: Image 6000659.jpg from the travel website folder, and on the right: Colour palette containing 5, 10 and 20 values.

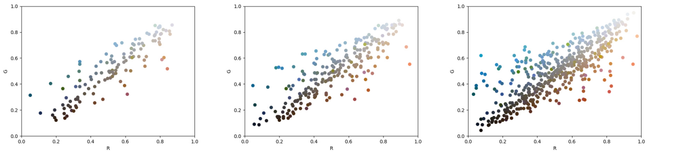

Fig. 4. Hotel: 451. On the left : Scatter Plot with 5 values extracted per image, in the middle: Scatter Plot with 10 values extracted per image, on the right side: Scatter Plot with 20 values extracted per image.

hindering the detection of nuances in images with analogous colour palettes.

In scatter plots with more extracted colour values, clustering or concentration in specific colour spectra is noticeable, implying the frequent occurrence or dominance of particular colour ranges. But these clusters can exacerbate the "curse of dimensionality", which shortens distances in higher-dimensional space, making it harder to distinguish similar images. In summary, the choice of the number of extracted colour values requires careful consideration to balance the distinguishability of similar images and the representativeness of the palette.

E. Visual Feature matching

To visually compare individual plots, a three-dimensional plot was generated using the colour palette of Hotel 451's `traffickcam` folder, with all data points in red. Subsequently, the `travel_website` folders of different hotels from the "5 hotels" subset were displayed in separate three-dimensional plots, using the colour green for better differentiation. To facilitate comparability, data points from both sets of plots were merged into a combined plot. This step aims to illustrate the visual differences and similarities between the colour palettes of the `traffickcam` and `travel_website` data sources.

The combined scatter plots reveal significant differences in the data-point distributions. Despite potential data quality limitations in the `traffickcam` data, a close correlation can be observed between the outer data points of the `travel_website` palette from hotel folder 451 and the matching plot of the same hotel folder. This is particularly evident in edge areas, where data points from hotel folder 451 lack nearby counterparts in the middle and right comparison plots. However, data points in the hotel folder show little distinctiveness in densely populated areas due to agglomeration. These suggest some consistency and correlation between the colour values in the different datasets.

F. Calculation of the Smallest Distance between Two Plots

Euclidean distance; a commonly adopted metric for evaluating colour similarity, is used to calculate the similarity between the colour vectors of the matching image and those in the image series. It measures the geometric distance between the colour points in three-dimensional space, where smaller distances indicate greater similarity, as expressed with the formula below.

$$Distance = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \quad (1)$$

As outlined above, certain frequently occurring colours may be poor identifiers, while rarer shades are more distinctive. The weighted distance could reduce clustering effects caused by a larger number of colour values, helping to distinguish between outer data points and potential false positives.

The Euclidean distance considers the coordinates (x, y, z) of the colour values in the three-colour channels (e.g., red, green, blue) for two data points. For each point in the matching image, it identifies the closest point from the respective hotel folder's image series. This calculation generates a numerical value that quantifies the similarity of the colour values. A lower value indicates higher similarity, while a higher value signifies a greater distance between the colour values.

V. Experimental Design

A. Dataset Configuration

In the course of the experiment, three additional subsets were added to the existing dataset "minimum requirements", to ensure that the results are robust and not limited to specific properties or patterns of a single dataset. Each of these subsets is made up of three individual datasets, each of which, in turn, comprises 100 different hotel folders.

The datasets, namely "unclean" dataset, "clean" dataset and "three beds" dataset, are explained in detail below. The "unclean" datasets include images from the `traffickcam` and `travel_website` folders. These folders were taken unchanged from the "minimum requirements" subset, resulting in significant variance in the number of images per folder. In total, the "three beds" datasets contain 6,562 images from `travel_website` and 1,281 from `traffickcam`.

The "clean" datasets were created through a step-by-step process. First, images from the `traffickcam` folders were removed. The remaining datasets were then cleaned by eliminating images taken outside hotels, images not clearly showing hotel rooms, and, as shown in Figure 6, pictures of individual objects or people.

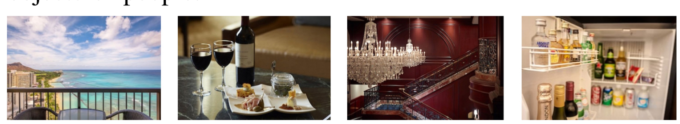

Fig. 6. Sample images that were removed for cleaning.

The remaining images from the `travel_website` data were further processed. In a "leave one out" procedure, an image from each hotel folder was selected for matching, ensuring that at least one other image in the folder showed the same setting. Ultimately, 5,384 images remained from the `travel_website` dataset, with 300 used for matching. This approach aimed to maintain the comparability and relevance of the selected images, considering the diversity of content.

The "three beds" datasets were developed from the "clean" datasets to optimise image matching. Four images were selected from each hotel folder in the `travel_website` data, exclusively depicting beds with similar colouring. Three images were retained in the folder, while the fourth was set aside for comparison, as shown in Figure 7.

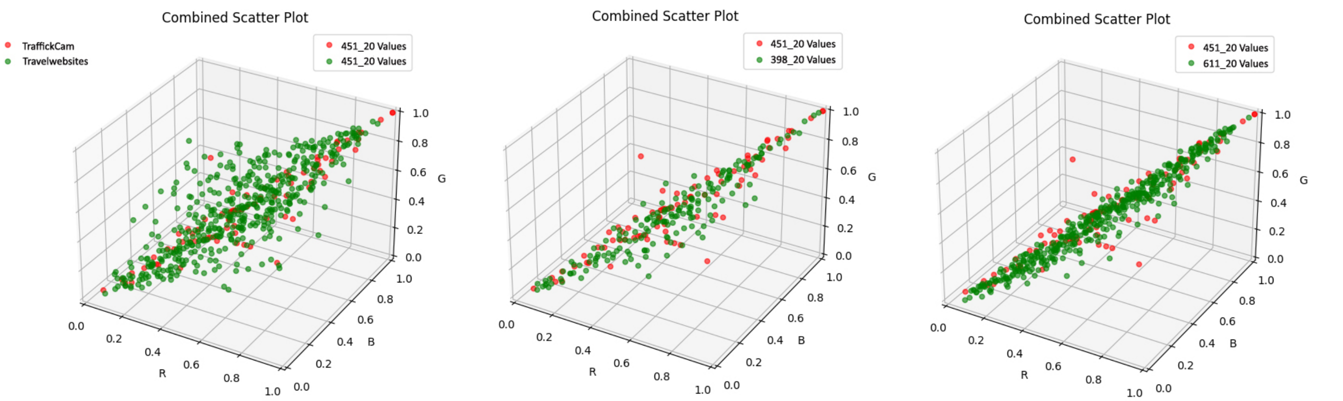

Fig. 5. Left side: Scatter Plot comparing colour values from hotel 451's photos from TraffickCam (red) vs. the travel website (green). Middle and right: Scatter Plot comparing colour values from hotel 451's photos from TraffickCam (red) vs. hotel 398's and hotel 611's photos from the travel website (green), respectively.

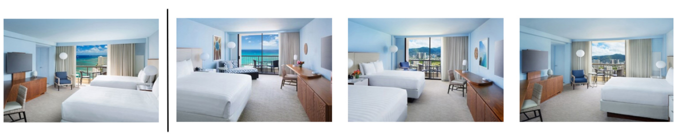

Fig. 7. Example from the three beds datasets. On the left side is the retrieval image, and on the right side the three images that remain in the Travel Websites folder.

B. Method Validation

The developed method is evaluated through supervised validation against known classes/labels. The primary objective is to verify the accuracy of the method in correctly matching images to the corresponding hotels. The validation process involves several steps. First, descriptors are created for each dataset, with colour palettes containing 1, 2, 5, 10, and 20 colour values stored in separate CSV files. Then, precise distance measurements between these colour palettes are conducted. To ensure the stability of the results, supervised validation is performed in three separate runs, with each run using different hotel folders to account for random variations in the test set. Performance is assessed using Top-K accuracy, providing an authentic evaluation of the test runs. The focus of this research is not on maximising precision, but on analysing the suitability of colour values as descriptors for a data triage.

VI. Results

For each test run, tables and diagrams were created to illustrate the achieved Top-K accuracies in percentages. The initial test run used the "unclean" datasets, with direct Euclidean distance as the comparison metric. These datasets included images from both the `travel_website` and `traffickcam` folders, with a minimum of 8 images in the `travel_website` folders and at least one in the `traffickcam` folder. The datasets reflected the original Hotels-50K dataset, including both hotel room images and unrelated scenes. A satisfactory Top-K accuracy was set at 95%, aligning with realistic thresholds. Threshold selection depends on the objective, with critical applications that require higher thresholds, while others may accept lower ones.

Extracted Values	Top-10	Top-20	Top-30	Top-40	Top-50	Top-60	Top-70	Top-80	Top-90	Top-100
1 Value	11,33	22,67	30,33	43,33	53,33	63,67	71,33	81,33	91,67	100,00
2 Values	10,33	22,67	32,67	44,33	52,33	62,00	74,00	84,00	91,67	100,00
5 Values	12,33	23,00	33,00	43,67	52,67	63,67	72,67	83,00	90,33	100,00
10 Values	12,33	22,00	33,00	42,67	51,67	62,67	71,33	82,33	90,00	100,00
20 Values	13,00	22,67	35,00	45,00	54,00	63,33	70,33	82,00	90,33	100,00

TABLE I
UNCLEAN DATASETS AVERAGE TOP-K ACCURACY EUCLIDEAN DISTANCES

Table I shows that the first test run with "unclean" datasets produced suboptimal results for all the extracted colours. Although the top 100 accuracy met the threshold 95%, the top 50 accuracy was approx. 50%, indicating that the matching image was found in the top 50% of hotel images in only half of the cases. This was largely due to the presence of out-of-scene images, whose colours often differed significantly from actual hotel room images. The blue spectrum, identified in previous studies as a high-quality identifier, contributed to false positives due to the inclusion of tourist-related images. These results confirmed the hypothesis. To resolve this, "clean" datasets were created by removing out-of-scene images.

Extracted Values	Top-10	Top-20	Top-30	Top-40	Top-50	Top-60	Top-70	Top-80	Top-90	Top-100
1 Value	42,00	62,00	72,67	81,00	89,33	92,67	95,67	97,33	99,33	100,00
2 Values	48,33	66,67	75,33	83,00	88,00	92,33	96,00	98,00	99,33	100,00
5 Values	44,33	59,67	68,00	78,00	85,00	90,33	94,67	96,67	99,00	100,00
10 Values	42,33	60,00	70,00	79,33	86,67	90,67	93,33	96,33	98,67	100,00
20 Values	43,33	57,33	69,33	80,67	87,00	92,00	96,00	97,33	99,00	100,00

TABLE II
CLEAN DATASETS AVERAGE TOP-K ACCURACY EUCLIDEAN DISTANCES

The "clean" datasets show significant improvement, as illustrated in Table II. In almost all cases, the 95% threshold was achieved for Top-70 accuracy. The best results were obtained using 2 and 20 descriptors, with both predicting with 96% reliability that the matching image is in the top 70%. Interestingly, 2 descriptors performed slightly better. When comparing 2 and 20 descriptors, it is possible that beyond a certain number of values, concentration centres develop, exacerbating the "curse of dimensionality." Using 2 descriptors may provide enough information for effective matching while minimising false positives, suggesting a balance between information retention and dimensionality. The superior performance of 2 descriptors compared to 5 may be due to more complex dynamics, as more descriptors do not always lead to better performance. However, a single descriptor per image appears insufficient.

Extracted Values	Top-10	Top-20	Top-30	Top-40	Top-50	Top-60	Top-70	Top-80	Top-90	Top-100
1 Value	51,00	68,33	80,33	88,00	92,33	95,33	96,67	99,00	99,67	100,00
2 Values	62,00	74,33	84,33	92,00	95,67	96,67	97,00	99,33	99,67	100,00
5 Values	55,33	66,33	77,00	83,00	89,67	93,33	96,00	98,33	99,33	100,00
10 Values	55,67	72,00	82,67	88,67	94,00	96,00	97,00	98,67	99,33	100,00
20 Values	53,33	70,33	80,33	86,67	90,67	94,33	96,00	98,67	99,67	100,00

TABLE III
RESULTS - THREE BEDS DATASETS AVERAGE TOP-K ACCURACY EUCLIDEAN DISTANCES

The results from the "three beds" datasets, as shown in Table III, confirm previous findings, with a top-50 accuracy exceeding 95% using 2 descriptors. Using 10 descriptors yielded slightly lower results, with a top-60 accuracy of 96%. Other descriptor values performed significantly worse, though they still exceeded the top-70 accuracy threshold, indicating acceptable outcomes. The consistent success of 2 descriptors suggests that more generalised descriptors may be better suited for this application. With the achieved Top-50 accuracy of more than 95% using two descriptors in the "three beds" dataset, the search field is significantly reduced. This data triage allows approximately half of the images to be excluded, thereby decreasing investigative effort and focussing on relevant matches within the remaining images.

A. Discussion

This study thoroughly investigated the effectiveness of colour-based descriptors for CBIR, analysing parameters such as the number of descriptors and distance metrics. The findings contribute to the development of advanced image search systems for the digital investigation of human trafficking, highlighting both strengths and limitations. The research identified the optimal number of descriptors, with 2 offering a balance between information content and dimension reduction, enabling efficient image retrieval with high Top-K accuracy. Furthermore, selective image sorting and dataset sampling improved the robustness of the method, reducing false positives and improving its practical applicability in various scenarios. The use of colour descriptors to narrow the search field enables efficient data triage, significantly reducing the number of images requiring analysis. This method can be integrated as a preliminary filter in forensic workflows, making the analysis more targeted and resource-efficient. Following this narrowing, texture-based methods could then be applied specifically to the most relevant images.

Figure 8 provides a clear distinction between the worst results of the "unclean" datasets, on the left, and the optimal results of the "three beds" datasets, on the right. The black diagonal line highlights the threshold that must be reached in order for the method to have a positive effect over random (i.e., result lines higher than this demonstrate the benefit of the proposed approach). The chart on the left of Figure 8 shows that even in the less successful test configurations, at least neutral results were achieved. On the other hand, the chart on the right (with the optimal results) illustrates the considerable added value that can be achieved not only by applying the proposed method, but also by cleaning up the datasets. Here, in particular, the study showed that the method is significantly influenced by the quality of the input images. Images with very similar colour spectra can affect the accuracy of the matching and lead to false positives. The same applies to scenes of foreign images, which have to be sorted out in advance.

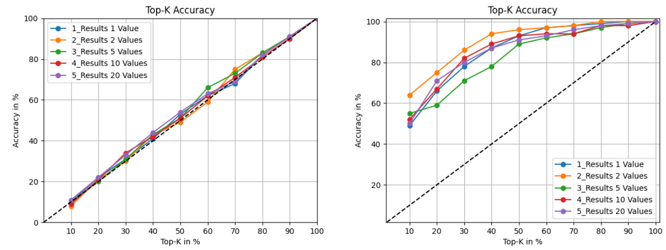

Fig. 8. Diagram on the left side: Top-K accuracy results unclean dataset 3, diagram on the right side: Top-K accuracy results for three beds dataset 2, both with the Euclidean distance

VII. CONCLUSION AND FUTURE WORK

This research demonstrates that the colour values of the images significantly contribute to the identification of hotel rooms in human trafficking and CSAM investigations. Our findings show that colour serves as a reliable descriptor under specific conditions, with promising applications for improving the quality and accuracy of digital investigations. Future work could focus on refining the quantisation process and exploring advanced colour extraction techniques to increase robustness. Although colour alone cannot fully address the complexities of contextual image analysis, continuous research such as [24] has demonstrated the value of combining colour with texture features. Addressing limitations of the Euclidean distance method, especially in "centres of concentration," along with histogram-based weighting, could further refine the colour-based analysis, reducing false positives when used in conjunction with texture data.

Furthermore, this colour-based approach could aid in filtering irrelevant images in CSAM, reducing the size of the dataset, and improving detailed analysis. Since hash databases are limited by their inability to identify dynamically generated images, minimal colour values and edge detection techniques may improve the classification and filtering of such files. In conclusion, colour descriptors offer a valuable and flexible tool for image recognition, search, and classification, capable of narrowing search fields and providing significant benefits in tackling complex investigative challenges.

REFERENCES

[1] "Regulation (COD) No 426/2022 of the European Parliament and of the Council of 19 December 2022 on

preventing and combating trafficking in human beings and protecting its victims," 2022, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX (Accessed on September 30, 2024).

[2] A. Stylianou, A. Norling-Ruggles, R. Souvenir, and R. Pless, "Indexing open imagery to create tools to fight sex trafficking," in *Proceedings of the IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, Washington, DC, USA, 2015, pp. 1–6.

[3] A. Stylianou, H. Xuan, M. Shende, J. Brandt, R. Souvenir, and R. Pless, "Hotels-50k: A global hotel recognition dataset," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 1, pp. 726–733, 2019.

[4] R. Kamath, G. Rolwes, S. Black, and A. Stylianou, "The 2021 Hotel-ID to Combat Human Trafficking Competition Dataset," *arXiv preprint arXiv:2106.05746*, 2021.

[5] X. Du, C. Hargreaves, J. Sheppard, F. Anda, A. Sayakkara, N.-A. Le-Khac, and M. Scanlon, "SoK: Exploring the State of the Art and the Future Potential of Artificial Intelligence in Digital Forensic Investigation," in *The 13th International Workshop on Digital Forensics (WSDF), held at the 15th International Conference on Availability, Reliability and Security (ARES)*. ACM, 08 2020.

[6] C. Hargreaves, F. Breitinger, L. Dowthwaite, H. Webb, and M. Scanlon, "DFPulse: The 2024 Digital Forensic Practitioner Survey," *Available at SSRN 4954821*, 2024.

[7] M. Scanlon, F. Breitinger, C. Hargreaves, J.-N. Hilgert, and J. Sheppard, "ChatGPT for digital forensic investigation: The good, the bad, and the unknown," *Forensic Science International: Digital Investigation*, vol. 46, p. 301609, 2023.

[8] European Parliament and Council of the European Union, "Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence," https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32024R1689, 2024, official Journal of the European Union, L 1689, published on July 12, 2024, accessed on November 2, 2024.

[9] F. Breitinger, J.-N. Hilgert, C. Hargreaves, J. Sheppard, R. Overdorf, and M. Scanlon, "DFRWS EU 10-Year Review and Future Directions in Digital Forensic Research," *Forensic Science International: Digital Investigation*, vol. 48, p. 301685, 03 2024.

[10] X. Du and M. Scanlon, "Methodology for the automated metadata-based classification of incriminating digital forensic artefacts," in *Proceedings of the 14th International Conference on Availability, Reliability and Security*, ser. ARES '19. New York, NY, USA: Association for Computing Machinery, 2019.

[11] O. Bamigbade, J. Sheppard, and M. Scanlon, "Computer Vision for Multimedia Geolocation in Human Trafficking Investigation: A Systematic Literature Review," *CoRR*, vol. abs/2402.15448, 2024.

[12] A. Stylianou, R. Souvenir, and R. Pless, "TraffickCam: Explainable Image Matching For Sex Trafficking Investigations," in *Proceedings of AAAI Fall Symposium Series (AAAI FSS-19): Artificial Intelligence in Government and Public Sector*, Arlington, Virginia, USA, 2019.

[13] H. Xuan, A. Stylianou, and R. Pless, "Improved embeddings with easy positive triplet mining," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, Snowmass, CO, USA, 2020, pp. 2463–2471.

[14] H. Xuan, A. Stylianou, X. Liu, and R. Pless, *Hard Negative Examples are Hard, but Useful*, ser. Computer Vision – ECCV 2020. Cham: Springer International Publishing, 2020, vol. 12359, pp. 126–142.

[15] B. Tseytlin and I. Makarov, "Hotel recognition via latent image embeddings," in *Advances in Computational Intelligence*, I. Rojas, G. Joya, and A. Català, Eds. Cham: Springer International Publishing, 2021, pp. 293–305.

[16] S. Black, A. Stylianou, R. Pless, and R. Souvenir, "Visualizing paired image similarity in transformer networks," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, January 2022, pp. 3164–3173.

[17] S. Saikia *et al.*, "Image feature representation using deep learning for instance search and scene recognition," Ph.D. dissertation, Universidad de León, León, Spain, 2021.

[18] N. A. M. Zin, R. Yusof, S. A. Lashari, A. Mustapha, N. Senan, and R. Ibrahim, "Content-based image retrieval in medical domain: A review," *Journal of Physics: Conference Series*, vol. 1019, no. 1, p. 012044, June 2018.

[19] R. Padhy, L. Samantaray, S. K. Dash, and J. Mishra, "Classification of high-resolution satellite image with content based image retrieval and local binary pattern," in *International Conference on Innovations in Intelligent Computing and Communications*. Springer, 2022, pp. 409–416.

[20] H. Cui, N. Giebink, J. Starr, D. Longert, B. Ford, and E. Léveillé-Bourret, "From noisy data to useful color palettes: One step in making biodiversity data fair," vol. 13971, 2023, pp. 469–481.

[21] "Framework (COD) No 106/2021 of the European Parliament with recommendations to the Commission of 20 October 2020 on ethical aspects of artificial intelligence, robotics and related technologies," 2020, https://oeil.secure.europarl.europa.eu/oeil/popups/printsummary.pdf (Accessed on September 30, 2024).

[22] H. Aboaisha, "The optimisation of elementary and integrative content-based image retrieval techniques," PhD thesis, University of Huddersfield, Huddersfield, United Kingdom, 2015.

[23] Y.-C. Hu, M.-G. Lee, and P. Tsai, "Colour palette generation schemes for colour image quantization," *The Imaging Science Journal*, vol. 57, no. 1, pp. 46–59, 2009.

[24] F. A. Alghamdi, "An effective hybrid framework based on combination of color and texture features for content-based image retrieval," *Arabian Journal for Science and Engineering*, vol. 49, no. 3, pp. 3575–3591, 2024.