

Filming the sound: Anomaly Detection on Audio Tape Recordings using Computer Vision Algorithms

Zafer Çınar^{1,†}, Alessandro Russo^{1,†}, Matteo Spanio^{1,*,†}, Niccolò Pretto^{2,†} and Sergio Canazza^{1,†}

¹Centro di Sonologia Computazionale (CSC), Department of Information Engineering, University of Padua, Via Giovanni Gradenigo, 6b, 35131, Padua, Italy

²Media Interaction Lab, Faculty of Engineering, Free University of Bozen-Bolzano, Via Bruno Buozzi, 1, 39100, Bozen, Italy

Abstract

The preservation of open-reel audio tapes is critical for maintaining valuable cultural and historical audio archives, yet current digitisation and analysis operations are often error-prone due to tape degradation and the long duration of the recordings. Considering the analog nature of this kind of recording, anomaly detection algorithms, applied to the video of the tape flowing on the playback head, can be used to detect errors and details with musicological value. This paper presents a new dataset of high-quality videos and a new algorithm for anomaly detection on audio tapes. Experimental results show notable improvements in detection performance, though false positives remain a challenge at higher speeds. Additionally, the new algorithm supports a wider range of playback speeds, improving its flexibility. This improvement is an important step towards a reliable implementation of the IEEE/MPAI CAE ARP standard (3302-2022).

Keywords

open reel audio tapes, irregularities detection, computer vision, preservation, restoration

1. Introduction

Nowadays, audiovisual archives are increasingly facing the challenge of preserving their collections from deterioration [1]. Digitization is a key solution, converting analog materials like photos, films, videos, and audio recordings into digital formats to mitigate physical degradation. However, the digitization process must be based on a scientific methodology to ensure minimal information loss. The Centro di Sonologia Computazionale (CSC) of the University of Padua has been working in audio document preservation for the last decade [2, 3], carrying out its research activity to develop a preservation methodology for audio documents [4]. At CSC, digitization goes beyond simply migrating the audio content; it includes gathering metadata and contextual information, such as photos and video documentation. This approach became essential when working with archives of electronic music composers like Luciano Berio and Luigi Nono, who left markings and notes on tapes, also as indications for live performances, when tapes were used almost like instruments on stage. Furthermore, the video documentation captures important information about tape conditions and mechanical issues that may affect its playback, such as dirt, loss of magnetic paste or deformations. Since reviewing hours of video can be time-consuming and prone to errors, artificial intelligence may assist archivists and

3rd Workshop on Artificial Intelligence for Cultural Heritage (AI4CH 2024, <https://ai4ch.di.unipd.it/>), co-located with the 23rd International Conference of the Italian Association for Artificial Intelligence (AIIA 2024). 26-28 November 2024, Bolzano, Italy

*Corresponding author.

†These authors contributed equally.

✉ zafer.cinar@studenti.unipd.it (Z. Çınar); alessandro.russo@dei.unipd.it (A. Russo); spanio@dei.unipd.it (M. Spanio); niccolo.pretto@unibz.it (N. Pretto); sergio.canazza@unipd.it (S. Canazza)

🌐 <https://matteospanio.github.io/> (M. Spanio);

<https://www.unibz.it/it/faculties/engineering/academic-staff/person/47860-niccolo-pretto> (N. Pretto);

<https://www.dei.unipd.it/~canazza/> (S. Canazza)

🆔 0000-0001-6691-759X (A. Russo); 0000-0002-2436-7208 (M. Spanio); 0000-0002-3742-7150 (N. Pretto); 0000-0001-7083-4615 (S. Canazza)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

researchers by automatically detecting points of interest on the surface of the tapes and, therefore, the corresponding moment on the digitized audio recording. The methodology proposed by CSC has been the core reference during the implementation of the IEEE/MPAI CAE ARP standard, approved in December of 2022¹. The implementation of a standard is often a long-term process that requires continuous updates and refinements to accommodate evolving needs and technological advancements. In the case of the IEEE/MPAI CAE ARP standard, the development of effective tools required several iterations of improvements since its first version [5]. As new technologies emerge, software must be re-evaluated and enhanced to address previous limitations, improve overall performance, and meet the evolving expectations of both the archival and technical communities. This is the case of the *Video Analyzer*, a component described in the standard developed to detect anomalies (also referred to as *irregularities*), such as splices, tape degradations or notes on the surface of the digitized tapes. This module analyzes a video framing a close-up of the tape flowing in front of the magnetic head during the digitization process. This paper contributes to this ongoing effort by presenting a series of significant improvements to the anomaly detection process in open-reel audio tapes. The key contributions of this work include:

1. a new dataset of high-quality videos can be used for testing new algorithms.
2. a new detection algorithm that shows a strong improvement in comparison to the initial algorithm implemented in the Video Analyzer component.
3. an extension of the supported playback speeds managed by the algorithm.

These improvements result in a more robust and accurate system for identifying irregularities on audio tapes, which will improve the reliability of the overall implementation of IEEE/MPAI CAE ARP standard and, therefore, foster correct preservation and reduce restoration and analysis efforts. The next Section provides a overall view on the background and related works at the base of this paper. Section 3 provides more information about the standard. Then, the proposed algorithm is described in Section 4. The experiment and dataset used to test the algorithm are reported in Section 5 along with its performance, compared to the previous version. Finally, Section 6 concludes the article with a discussion of the results and further opportunities for development.

2. Related work

Anomaly detection has been a lasting yet active research area in various research communities for several decades [6]. Traditional methods often rely on machine learning models, such as convolutional neural networks (CNNs) and support vector machines (SVMs), to recognize patterns in images or video frames that deviate from normal behaviour. Studies on the application of deep learning models to anomaly detection are still ongoing, with efforts concentrated on capturing high-dimensional representations of normal data [7]. More recently, unsupervised approaches, including autoencoders and Generative Adversarial Networks (GANs), have emerged as effective alternatives for detecting abnormalities, particularly in cases where labeled data is limited [8, 9]. However, in this specific context, using deep learning models poses several challenges. One major issue is the need for large-scale datasets, which are difficult to obtain due to the time-consuming and error-prone nature of annotating frames for tape anomalies. The lack of sufficient labeled data makes it impractical to train deep learning models effectively. In the specific domain of anomaly detection in time-based media, like videos, frame-by-frame comparisons are usually adopted to detect subtle visual differences [10]. The process typically involves analyzing pixel-level changes and identifying unusual patterns such as signal damages or inconsistencies. For example, background subtraction methods have been used in video surveillance to detect anomalies, though they suffer from false positives when exposed to subtle variations like lighting changes [11]. These techniques have also been adopted for other applications, such as medical image analysis and industrial monitoring [12, 13].

¹Link to the standard <https://standards.ieee.org/ieee/3302/11006/> (Last Accessed: November 4, 2024)

2.1. Overview of the Video Analyzer

The Video Analyzer is one of the main components of the IEEE/MPAI CAE ARP standard (an overall description of the standard will be provided in Section 3). This component implements an anomaly detection algorithm on the video of the tape flowing on the playback head of the tape recorder. The frame of each irregularity identified by the algorithm and related metadata are the output of this component. The first version of the Video Analyzer featured an innovative method for detecting irregularities on the surface of open-reel audio tapes by analyzing videos produced during the digitization process [14]. This approach enabled a detailed frame-by-frame inspection, allowing for the identification of physical issues such as splices, scratches, and deformations. The system employed advanced computer vision algorithms, notably the Generalized Hough Transform and Speeded Up Robust Features (SURF), to identify regions of interest (ROIs) on the tapes and detect potential anomalies. The detection process involved several key steps. First, the system focused on detecting ROIs in the section of the tape beneath the reading head, using fixed elements like the pinch roller, a rotating rubber wheel which accompanies and controls the movement of the tape, as reference points to ensure consistency across frames. In the next phase, consecutive video frames were compared to identify significant differences. This was done through a pixel-level analysis, where changes in pixel intensity were monitored. If the number of differing pixels between two frames exceeded a set threshold, the system flagged the frame as containing an anomaly. The system generated a difference image for each pair of frames, representing potential irregularities in the tape. After the anomaly detection phase, the identified irregularities were not classified immediately. Instead, the output from this process was passed to the Tape Irregularity Classifier, a module within the IEEE/MPAI CAE ARP standard. At this stage, a convolutional neural network was used to categorize the detected irregularities, identifying specific issues such as splices, dirt, or other forms of damage on the tape's surface. One of the primary limitations of the original methodology was its reliance on the PAL (720x576) video format, which was the standard for most of the archived video recordings. The use of PAL video introduced several challenges. For instance, the low resolution and interlaced nature of PAL video (25 interlaced frames per second) affected the accuracy of anomaly detection, particularly during the image classification stage. The misalignment between odd and even lines due to interlacing reduced the precision of the convolutional neural network in identifying anomalies, as it disrupted the visual clarity required to detect subtle tape irregularities [14]. To overcome these issues, high-definition (HD) video is necessary. By increasing the video resolution up to Full HD (1920x1080) and frame rate to 50 fps (progressive), the enhanced method aims at reducing motion blur by using a shutter speed of 1/100 of a second and improving the detail captured in each frame. This allows for a more precise identification of anomalies, such as small scratches or splices that may go unnoticed in lower-quality videos. This improvement also allows for an overall better performance of machine learning models, which can leverage the additional details to detect more complex irregularities. Thus, the transition to FHD video is crucial to refining the anomaly detection process and achieving more accurate and reliable results.

3. IEEE/MPAI CAE ARP

MPAI is an international, independent, non-profit organization dedicated to developing standards for AI-based data coding. MPAI Context-based Audio Enhancement (MPAI-CAE) standard aims to enhance the user experience in audio-related applications such as entertainment, communication, teleconferencing, gaming, post-production, and restoration. The MPAI-CAE international standard was approved in May 2022 and subsequently adopted by the IEEE Standards Association as 3302-2022 in December of the same year. The presented work falls within the Audio Recording Preservation (ARP) use case, a part of the MPAI-CAE standard. The IEEE/MPAI CAE ARP standard provides precise software references for audio document preservation. Its technical specifications adopt the preservation methodology developed at the CSC, incorporating AI-based computational tools to extract information from digitized audio/video of analog open reel tapes. The use of AI enables the automatic detection of irregularities on the surface of the tape, improving precision and speed in selecting and extracting irregularities,

such as splices, marks, and loss of magnetic paste. The technical architecture of the standard includes five modules that target and process different digital inputs: the Audio Analyzer, the Video Analyzer, the Tape Irregularity Classifier, the Tape Audio Restoration, and the Packager. The initial version of the Video Analyzer was implemented by training the system using a dataset of videos created by the CSC during numerous digitization projects. These videos were recorded during the A/D transfer of the signal to provide a visual record of magnetic tapes, most of which were preparatory tapes containing electronic music recordings. Documenting the presence of splices, different tape segments, possible alterations, annotations, and marks added by the composer can be extremely useful for reconstructing the philological history of a given work as well as detecting parts that could require an audio restoration [15]. Moreover, the video provides valuable information regarding the preservation conditions of audio documents, allowing to keep trace of splices, marks, and other surface irregularities.

4. Algorithm Description

The proposed method maintains the same framework as the existing approach, where consecutive frame pairs are compared as the video plays. The logic behind this method is to treat the problem as motion detection. While there is constant movement in the Region of Interest (ROI), areas without irregularities appear static, like still images. This allows for the use of a traditional motion detection algorithm, with irregularities seen as moving objects and regular regions as the background. The modification introduced in this work is based on frame differencing with additional filtering steps. This approach is chosen to improve the accuracy of detecting irregularities by focusing on meaningful changes between consecutive frames while reducing the impact of noise and irrelevant variations. By adding filtering, the method improves at distinguishing real irregularities from minor fluctuations that don't indicate actual issues. The flowchart in Figure 1 shows the frame differencing process with additional steps that incorporate Otsu's method [16] for thresholding and determining irregularities.

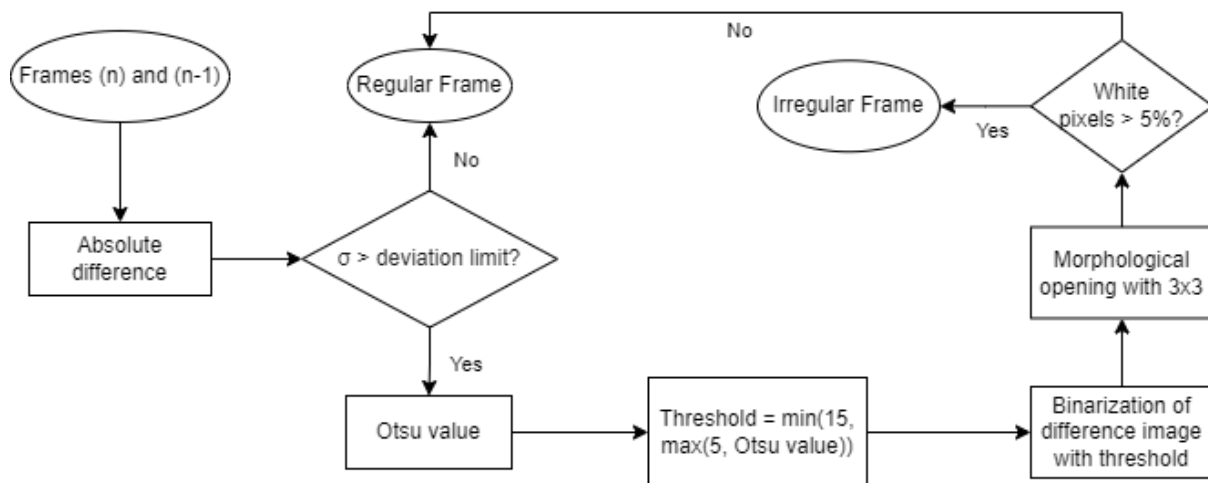


Figure 1: Flowchart of the frame differencing method with additional filtering.

The first step in this method is identifying ROI, based on the latest techniques: where the detection of the reading head is carried out using the Generalized Hough Transform and Speeded-Up Robust Features (SURF) is used for detecting the position of the pinch roller. In this implementation, the ROI of the tape area is reduced by half to decrease the computational load and improve accuracy. While this method reduces computational overhead, it may increase the number of irregularities detected in the case of problems that extend over long portions of the tape (for example, writing on the tape or a large scratch) or tapes running at very low speeds. While this solution creates a certain amount of “noise”, later modules of the standard are able to handle the extra frames, and this approach also reduces information loss. This step helps the method focus only on the relevant parts of the video,

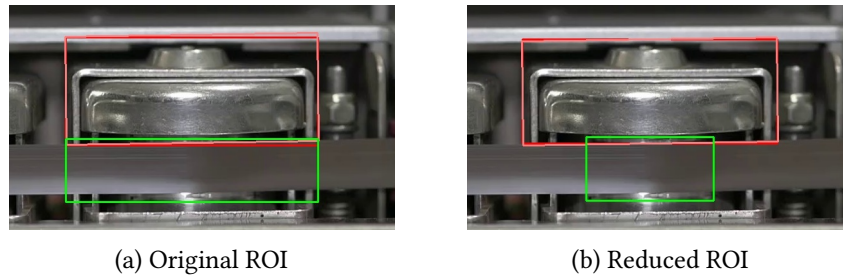


Figure 2: Comparison of original and new Regions of Interest (ROIs). The green rectangles highlight the ROIs while red rectangles highlight the reading heads.

making it more efficient and accurate. Figure 2b shows the reduced ROI.

Next, the method calculates the absolute difference between pairs of consecutive frames to create a difference image. This image shows the intensity variations between frames, with significant changes indicating possible irregularities. By focusing on these intensity differences instead of just color changes, the method can better detect motion and irregularities, ensuring that subtle but important differences are not missed. This approach treats differences more selectively, unlike the previous method, which treated all differences the same. Figure 3c shows the result of the absolute difference between the two consecutive frames Figure 3a and Figure 3b.

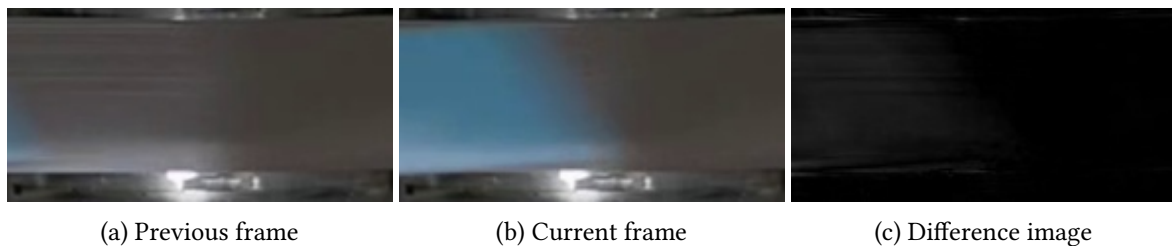


Figure 3: Comparison of previous and current frames, with the calculated difference image.

After creating the difference image, the standard deviation is calculated to decide if a frame needs further evaluation. For a frame to be evaluated, the standard deviation must exceed the threshold referred to as deviation limit in Figure 1. Frames with a standard deviation below this limit are considered regular and are excluded from further processing. This step reduces the computational load by ensuring that only frames with significant intensity variations are processed. The deviation limit varies based on the tape speed, with values set as follows: 2.25 for 30 inches per second (ips), 2.5 for 15 ips, 2.6 for 7.5 ips, and 2.75 for 3.75 ips. This separation is made to address the intensity change with different speeds, as higher speeds result in greater motion blur and reduced intensity. For frames that pass the standard deviation check, Otsu's method is used to find the right threshold for binarizing the difference image. Depending on the result, either a global or Otsu's threshold is selected to create a binary motion image. To improve the accuracy of detecting irregularities, upper and lower limits are set on the threshold values. The upper limit is 15 to prevent the threshold from being too high and missing parts of an irregularity by potentially filtering out portions of it. The lower limit is 5 to stop the threshold from being too low, which could cause false positives by highlighting insignificant portions of the frame. This approach adapts the thresholding process for each frame and makes irregularity detection more robust and reliable. Once the threshold is set, it is applied to the difference image to create a binary motion image. This image shows potential irregularities as white pixels on a black background, making it easier to identify significant differences between frames. The binary motion image serves as a key intermediate step. Figure 4a shows the binarized image obtained by thresholding the difference image in Figure 3c. To further improve the binary motion image, an opening operation is applied with a kernel 3x3. This process, which involves erosion followed by dilation, removes small, irrelevant artifacts caused by tape vibration or other external factors while keeping the main irregularities intact. This step

improves the clarity of potential irregularities, making them easier to detect and analyze in the final evaluation phase. Finally, the method counts the number of white pixels in the processed image and compares it to a set threshold of 5% of the total pixels to determine if the frame contains an irregularity. If the count exceeds this threshold, the frame is marked as irregular; otherwise, it is classified as regular. Figure 4b shows the image after morphological opening. Figure 5 and Figure 6 illustrate the processes applied to an annotation and a shadow, respectively.

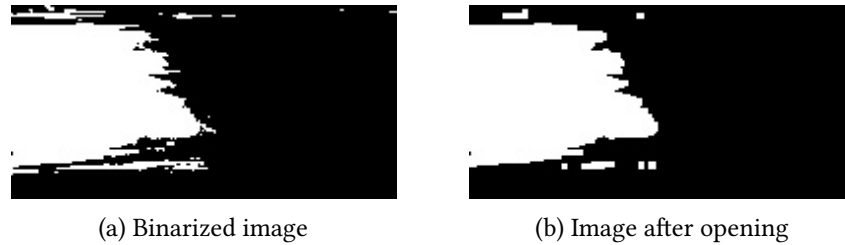


Figure 4: Progression from the difference image to thresholding, followed by morphological opening.

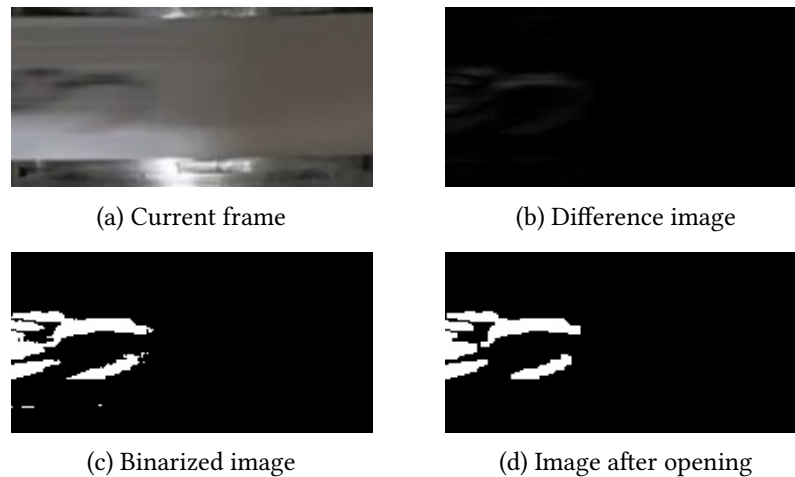


Figure 5: Progression from the current frame to final image for an annotation.

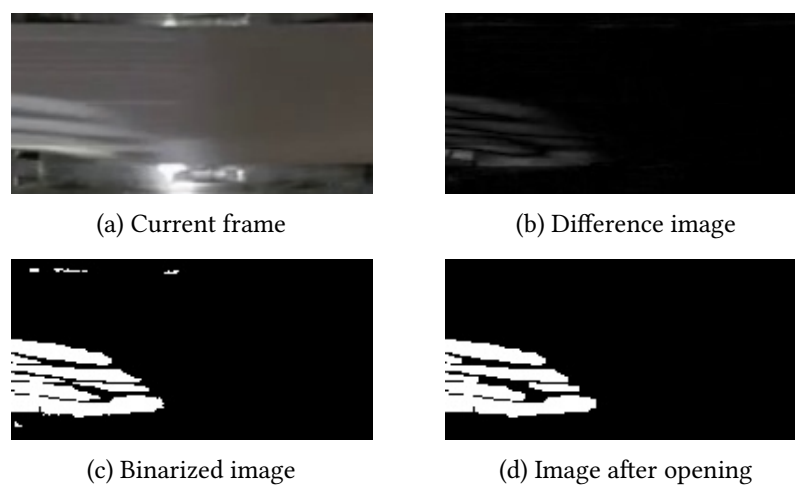


Figure 6: Progression from the current frame to final image for a shadow.

A further improvement concerns the quality of the videos provided as input of the algorithm. The analysis was previously conducted on videos in PAL format at 25 fps interlaced with a resolution of

720x576. The new algorithm works on high-definition videos with 50 fps (progressive), a fixed shutter speed of 1/100 of a second and a resolution of 1920x1080. This enhancement in video quality allows for more precise detection of irregularities, as finer details and smaller anomalies can be captured and analyzed more effectively. Since the video documentation began before the development of this project, many videos were of lower quality and interlaced. One approach to handle the misalignment caused by interlacing was to separate the even and odd fields, which mitigated the misalignment but reduced the resolution to 720x228. To make effective use of these existing resources, the new method provides a solution by employing interpolation-based deinterlacing. This approach ensures that the captured irregularities are not limited to the reduced resolution of 720x228 but are instead maintained at the original full resolution. By preserving the original dimensions, the method aims to enhance the accuracy of irregularity detection and also addresses the challenge of misaligned lines, which could impact classification in the future. Figure 7a shows a frame from the original interlaced video and Figure 7b shows the resulting image of deinterlacing.

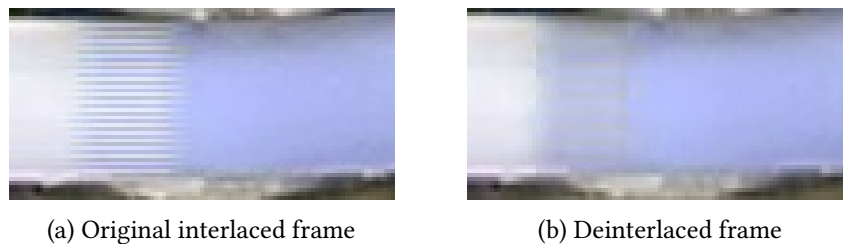


Figure 7: Comparison between an interlaced frame with misaligned lines and the deinterlaced frame.

In this novel version, the speed options have been expanded. The previous version of the Video Analyzer was specifically calibrated for detecting the anomalies on tapes recorded at 7.5 ips and 15 ips. However, the new updated version also supports lower and higher speeds (3.75 ips and 30 ips). This change makes the method more flexible and adaptable, allowing it to handle a wider range of tape playback scenarios.

5. Experiment and evaluation

To evaluate the performance of the improved anomaly detection method, an experiment was conducted using videos of open-reel tapes featuring various manually annotated irregularities (mainly splices, as they are the most common one). The test aimed at comparing the improved detection method against the original one across different playback speeds, providing a comprehensive analysis of precision and recall. The videos were captured at four distinct speeds: 3.75 ips, 7.5 ips, 15 ips, and 30 ips, with tape durations varying accordingly—10 minutes for 3.75 ips, 34 minutes for 7.5 ips, 10 minutes for 15 ips, and 9 minutes for 30 ips. At 3.75 ips, there were 14 splices present. At 7.5 ips, the tape had 66 splices, along with 4 annotations, 3 shadows, and 3 end-of-tape markers (a full description of this kind of irregularities can be found in [4]). For 15 ips, the tape featured 55 splices and 1 shadow. At 30 ips, the tape had 93 splices, 2 end-of-tape markers and 2 annotations. For the sake of comparison completeness, the authors also modified the previous version of the algorithm to support two additional speeds. The experiment focused on two key metrics: precision (the ability to correctly identify anomalies without generating false positives) and recall (the ability to detect all present anomalies). Both metrics were measured for each playback speed to compare the results of the old and new methods. The results, summarized in Table 1, show a notable difference between the two approaches. The new method successfully detected all irregularities at every speed but produced some false positives, while the old method demonstrated inconsistent performance, particularly struggling at lower speeds.

The new method detected all irregularities across the given speeds. However, in some cases, it generated some false positives. This was especially noticeable at 30 ips, where the motion blur caused by the higher speed required an increment of the sensitivity value for detecting the frame differences,

Table 1
Evaluation for different speed options

Tape Speed	Precision - New	Recall - New	Precision - Old	Recall - Old
3.75ips	0.9333	1.0000	0.0000	0.0000
7.5ips	0.9268	1.0000	0.4156	0.8421
15ips	0.9655	1.0000	0.9643	0.4821
30ips	0.8818	1.0000	0.9655	0.8660

which leads to more false positives under the presence of lighting changes and vibrations. It should be noted that some of the false positives are duplicates of the same irregularities. The method was designed to ensure that closely occurring irregularities are not missed, which sometimes results in duplicates. In future releases, the method can be improved by merging consecutive irregularities into one after the classification. This change could help in reducing false positives and enhancing the overall detection precision. The old method faced significant challenges, specifically at 3.75 ips, where it failed to detect any irregularity. This shortcoming is likely due to its reliance on detecting large pixel differences between frames, which are less noticeable at slower speeds. Consequently, the old method showed some improvements at the highest speed (30 ips), where pixel variations are more pronounced, with a lower detection precision in comparison to the new improvement. Additionally, due to its reliance on pixel count, the old method struggled to detect irregularities other than splices. This limitation arose because irregularities such as shadows and annotations often affect only a smaller area of the tape, making them less noticeable in comparison with splices. The new method addresses these issues more effectively: by focusing on intensity differences between frames, and using filtering techniques, it provides a more accurate detection. As a result, the overall performance of the novel method is better than the previous one, providing a more accurate and consistent irregularity detection across different speeds and various video conditions.

6. Conclusions

This work introduced significant improvements in the automatic detection of superficial irregularities on open-reel audio tapes using computer vision techniques. The key contributions include the development of a new dataset of high-quality videos², an enhanced detection algorithm with improved accuracy, and an expanded range of supported playback speeds. Despite these advancements, the system still faces certain limitations. The increased recall of the new algorithm, especially at higher playback speeds, led to false positives, particularly due to lighting changes and vibrations. Additionally, double detections of the same irregularities highlighted the need for further refinements, such as post-processing techniques that can merge closely occurring irregularities into a single detection. Future works will focus on addressing these limitations. Possible improvements will include refining the filtering process to reduce false positives and optimizing the classification module to handle more complex irregularities. Furthermore, expanding the algorithm's capabilities to work with other tape formats and video resolutions could enhance its applicability to a broader range of archival materials, supporting better preservation and restoration efforts in audiovisual archives.

Acknowledgments

This work is partially supported by the SYCURI Project, funded by the University of Padova in the Program “*World Class Research Infrastructure*”.

²The dataset presented in this article is available on Zenodo. The assigned DOI is 10.5281/zenodo.14028922. Please refer to this repository to access data and additional details.

References

- [1] F. Rumsey, Will you be mine forever? Audio archiving, multitracks, and 90s digital, *Journal of the Audio Engineering Society* 68 (2020) 304–307. URL: <https://aes2.org/publications/elibrary-page/?id=20736>.
- [2] S. Canazza, G. De Poli, Four decades of music research, creation, and education at Padua's Centro di Sonologia Computazionale, *Computer Music Journal* 43 (2020) 58–80. doi:10.1162/comj_a_00537.
- [3] S. Canazza, G. De Poli, A. Vidolin, Gesture, music and computer: The Centro di Sonologia Computazionale at Padova University, a 50-year history, *Sensors* 22 (2022). doi:10.3390/s22093465.
- [4] N. Pretto, C. Fantozzi, E. Micheloni, V. Burini, S. Canazza, Computing methodologies supporting the preservation of electroacoustic music from analog magnetic tape, *Computer Music Journal* 42 (2018) 59 – 74. doi:10.1162/comj_a_00487.
- [5] M. Bosi, N. Pretto, M. Guarise, S. Canazza, Sound and music computing using AI: Designing a standard, in: *Proceedings of the 18th Sound and Music Computing Conference, Virtual, 2021*, p. 215 – 218. doi:10.5281/zenodo.5045003.
- [6] G. Pang, C. Shen, L. Cao, A. V. D. Hengel, Deep learning for anomaly detection: A review, *ACM Computing Surveys* 54 (2021). doi:10.1145/3439950.
- [7] R. Chalapathy, S. Chawla, Deep learning for anomaly detection: A survey, *ArXiv abs/1901.03407* (2019). URL: <https://api.semanticscholar.org/CorpusID:57825713>.
- [8] T. Schlegl, P. Seeböck, S. M. Waldstein, U. Schmidt-Erfurth, G. Langs, Unsupervised anomaly detection with generative adversarial networks to guide marker discovery, in: M. Niethammer, M. Styner, S. Aylward, H. Zhu, I. Oguz, P.-T. Yap, D. Shen (Eds.), *Information Processing in Medical Imaging*, Springer International Publishing, Cham, 2017, p. 146 – 157. doi:10.1007/978-3-319-59050-9_12.
- [9] X. Xia, X. Pan, N. Li, X. He, L. Ma, X. Zhang, N. Ding, GAN-based anomaly detection: A review, *Neurocomputing* 493 (2022) 497–535. doi:10.1016/j.neucom.2021.12.093.
- [10] W. Sultani, C. Chen, M. Shah, Real-world anomaly detection in surveillance videos, in: *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 6479–6488. doi:10.1109/CVPR.2018.00678.
- [11] S. Qasim, K. N. Khan, M. Yu, M. S. Khan, Performance evaluation of background subtraction techniques for video frames, in: *2021 International Conference on Artificial Intelligence (ICAI)*, 2021, pp. 102–107. doi:10.1109/ICAI52203.2021.9445253.
- [12] L. Nans, C. Mediavilla, D. Marez, S. Parameswaran, Leveraging motion saliency via frame differencing for enhanced object detection in videos, in: M. S. Alam, V. K. Asari (Eds.), *Pattern Recognition and Tracking XXXIV*, volume 12527, International Society for Optics and Photonics, SPIE, 2023, p. 125270V. doi:10.1117/12.2678373.
- [13] A. Wang, Y. Ji, M. Chen, Y. Liu, Z. Li, S. Yan, Moving object recognition on production line based on adaptive frame differencing algorithm, in: *2024 36th Chinese Control and Decision Conference (CCDC)*, 2024, pp. 966–971. doi:10.1109/CCDC62350.2024.10587928.
- [14] A. Russo, M. Spanio, S. Canazza, Enhancing preservation and restoration of open reel audio tapes through computer vision, in: G. L. Foresti, A. Fusiello, E. Hancock (Eds.), *Image Analysis and Processing - ICIAP 2023 Workshops*, Springer Nature Switzerland, Cham, 2024, pp. 297–308. doi:10.1007/978-3-031-51026-7_26.
- [15] N. Pretto, E. Micheloni, A. Chmiel, N. D. Pozza, D. Marinello, E. Schubert, S. Canazza, Multimedia archives: New digital filters to correct equalization errors on digitized audio tapes, *Advances in Multimedia* (2021). doi:10.1155/2021/5410218.
- [16] N. Otsu, A threshold selection method from gray-level histograms, *IEEE Transactions on Systems, Man, and Cybernetics* 9 (1979) 62 – 66. doi:10.1109/TSMC.1979.4310076.