



HAL
open science

Watching Historical Films Through AI: Reflections on Image Retrieval from Heritage Collections

Beatriz Tadeo Fuica, Olivier Buisson, Claude Mussou

► **To cite this version:**

Beatriz Tadeo Fuica, Olivier Buisson, Claude Mussou. Watching Historical Films Through AI: Reflections on Image Retrieval from Heritage Collections. *Cinergie – Il cinema e le altre arti*, 2021, 10.6092/issn.2280-9481/13082 . hal-03975105

HAL Id: hal-03975105

<https://hal.science/hal-03975105v1>

Submitted on 6 Feb 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Watching Historical Films Through AI: Reflections on Image Retrieval from Heritage Collections

Beatriz Tadeo Fuica*^a

Olivier Buisson^b

Claude Mussou^b

^a Université Sorbonne Nouvelle, Paris 3 (France)

^b French Audiovisual National institute (INA) (France)

Submitted: June 16, 2021 – Revised version: October 21, 2021

Accepted: October 23, 2021 – Published: December 20, 2021

Abstract

Computer tools allow us to watch films differently and offer film scholars the opportunity to ask diverse research questions and imagine innovative methods. This article focuses, more precisely, on how film historians could benefit from Artificial Intelligence (AI), a tool which has not been used much in the field. The objective is to share a first round of experiments using Snoop, an AI developed by a collaboration between the French National Audiovisual Institute (INA) and the National Institute for Research in Digital Science and Technology (INRIA), on a corpus of historical films that, because they have been aired on television, are available for research from the INA collections. Following an exploratory approach, Snoop allowed us to retrieve shared motifs in a corpus of historical films. In addition to discussing challenges and findings, this article aims to evaluate the overall experience and open new lines of inquiry. From a historiographical point of view, we aim to assess whether, and if so, how, AI systems are useful tools for the implementation of different methods that both answer new research questions and revisit old ones.

Keywords: Film History; Audiovisual Archives; Digital Cinema; Artificial Intelligence; Object Retrieval.

* ✉ beatriz.tadeo-fuica@sorbonne-nouvelle.fr

1 Introduction

In the aftermath of the Second World War, numerous Cinematheques started opening around the world and the activities of the International Federation of Film Archives (FIAF) regained force, facilitating an exhibition circuit via which the public regained contact with archival films (Tadeo Fuica 2019: 28–32). This contributed to questioning the ways in which film history had been written until that moment and gave birth to a deep historiographic debate. In response to histories such as those written, for example, by Maurice Bardèche and Robert Brasillach (1935), which were mostly based on cinephiles' memories, authors such as George Sadoul (1946, 1947, 1948, 1949) and Jean Mitry (1968) developed more rigorous methodologies, highlighting the importance of film watching and contextualisation (Louis 2020: 117–30). A few decades later, the combination of archivists' needs to preserve their materials and the academic willingness to regain contact with historical approaches, prepared the terrain for a new turn in the field (Elsaesser 2012: 592–93). The 1978 Brighton FIAF Congress is broadly considered as the event which materialised this change because it gave scholars the opportunity to watch archival pieces of early cinema (Gaudreault et al. 2012: 3). This contact with archival treasures allowed researchers to revisit the teleological approaches followed until then, which had marginalised and undervalued early cinema (Gaudreault and Gunning 1989). This congress also served to emphasise the need for collaboration between archivists and historians to move the discipline forward (Gaudreault 2006, Gunning 2006).

Nowadays, artificial intelligence (AI), defined as a system that is able to “correctly interpret external data, learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation” (Kaplan and Haenlein 2019: 17), allows us to watch films differently. The information provided by AI-based data gathering and analysis offers an opportunity to ask new questions and imagine innovative research methods that include not only working with a larger number of films, but also interact with the machine to retrieve images through suggestions made by the researcher or the machine itself. These can be then analysed within a specific corpus, moving away from previous studies that focused on the narrative of a single film title or the comparison of a few examples. However, film historians have not yet greatly benefited from the possibilities provided by AI. When planning this article, we wondered if this was due to technical limitations, the fact that this technology is not really helpful to respond to the kind of research questions proposed by film historians (or film scholars, more broadly), or a combination of both. In seeking to find the answers to these questions, this article discusses a first round of experiments using Snoop, an AI developed by a collaboration between the French National Audiovisual Institute (INA) and the National Institute for Research in Digital Science and Technology (INRIA), on a corpus of historical films that, because they have been aired on television, are available for research from the INA collections.¹ Our objective is to evaluate this experience and present new possibilities for the use of AI in film history scholarship. Above all, we aim to assess whether, and if so, how, AI can lead to develop innovative methods to answer new research questions and revisit old ones.

2 Exploring the Use of AI for Film History

Almost a decade ago, the 2012 special issue of the journal *Frames*, edited by Catherine Grant, demonstrated that using digital tools in film and moving image studies was not an exceptional practice. As time passes, the studies of the crossover between Digital Humanities and Film and Media Studies are gaining momentum, and several scholars are reflecting on disciplinary boundaries; on collaborations between the humanities, archives and computer scientists; and on the exciting research paths that this interdisciplinary approach might open up (Burghardt et al. 2020, Heftberger 2018, Latsis and Ingravalle 2017). As an example, in 2020, the *Quarterly for Digital Humanities* devoted a special issue to the subject of Digital Humanities and Film Studies and provided a thorough state-of-the-art assessment of the many tools and methods that have been explored by scholars until now. Among the methods presented in the introduction to this special issue, the editors review the origins of Multimedia Information Retrieval (MMIR – the field associated with automatic film analysis in computer science) and trace its development in the last twenty years while pointing at diverse applications (Burghardt

1. The INA is responsible for the legal deposit of broadcast material in France. It records and documents the programs aired around the clock by over 100 TV stations. Over 22 million hours of broadcast material are available for research purposes, among which films that have been broadcast on TV.

et al. 2020). They also focus on different branches of MMIR such as Content-based image retrieval (CBIR), Content-based video retrieval (CBVR) and Content-based audio retrieval (CBAR) to highlight how these methods have recently improved and the fact that “MMIR has reached a level of maturity that can provide a real added value for film scholars and the digital humanities in general” (Burghardt et al. 2020: 4).²

The experiments and reflections presented in our article are based on MMIR using deep learning technology. Our AI, Snoop, was set to deliver large groups of still frames showing the images of similar specific objects found in different films of a corpus. From a methodological point of view, we wondered if we would be really studying films. After all, we were going to reduce complex narratives to pages and pages of still images. However, we focused on the fact that using an AI to retrieve images was going to allow us to see what the naked eye would not be able to find. In addition, this methodology reminded us of the relationship that photography and cinema have had since the beginning, and which has not disappeared with the digital turn (Mulvey 2006, Rombes 2017). When planning the experiments, we also reflected upon why we wanted to retrieve images of specific objects from films and to what extent knowing about the presence of certain objects would open paths for innovative ways of analysing films or complementing current studies that do not rely on these tools. We wondered whether, for instance, it would be important, for a film historian to know how many cars were featured in a specific film or group of films. The answer to this question is likely to depend on the purpose of the specific research project. In our case, counting the occurrences of similar objects within our corpus – or even in one specific film from within it – was not our main point of interest. We took an exploratory approach in which we privileged discovering and finding content overlaps between the films, even if only one occurrence was present in one of them, knowing that Snoop, at this stage, was likely to miss some reoccurrences of similar objects. We were more interested in establishing comparisons to develop qualitative rather than quantitative analyses.

Our work is situated within the framework of “distant viewing” and “deep watching” as stipulated by Arnold and Tilton (2019), and Bermeitinger et al. (2019), respectively. At the same time, our aim diverges from theirs. On the one hand, Arnold and Tilton utilised their toolkit in order to compare the role and presence of the female protagonists of two American situational comedies – *Bewitched* (1964–1972) and *I Dream of Jeannie* (1965–1970) – based on the facial detection of the main actors and actresses (Arnold and Tilton 2019: 6–7), and the analysis and exploration of stylistic features in a collection of documentary photographs (Arnold and Tilton 2019: 7–10). On the other hand, Bermeitinger et al. introduced deep watching by focusing on the recognition of a specific set of political symbols in a carefully chosen corpus of audiovisual material depicting the nationalist leader Stepan Bandera (Bermeitinger et al. 2019). Although the theoretical approaches of both research groups are in practice sufficiently close to ours, they conducted more highly controlled experiments than us and the results they obtained combine qualitative and quantitative analysis in ways that our experiments do not. As explained below, at this stage, the evaluation of our experimental conditions and the research paths we intended to explore, made us focus only on qualitative analyses.

We have applied Snoop to a corpus composed of sixty-four archival films made between the years 1902 and 1952. Arnold et al. emphasise that analysing archival films with computer-based algorithms presents limitations since “[t]here is no guarantee that they will perform well on historic materials (black and white, grainy film stock, previously digitized at a low resolution)” and warn that “[e]xtensive testing and tuning is often required to achieve acceptable results” (Arnold et al. 2019: 5). However, these barriers did not obstruct our main aim, which is to reflect on a first set of exploratory experiences. From the beginning, we were aware that, like other AI systems, Snoop could miss many objects. The accuracy of AI’s results depends on training. Precision increases when more examples are loaded onto the AI, and further experiments are refined and repeated. The research presented in this article was based on a single round of experiments and this was one of the reasons for deciding not to provide a quantitative analysis. Although the precise knowledge of how many cars appeared in our selection was not a significant focus of our study, finding out that crowds, uniformed men and close-ups of hands were present in large quantities within our corpus triggered a kind of comparative analysis that has not been carried out until now. The images retrieved by the AI allowed us to elaborate certain hypotheses without having the need to watch all the films, trust our memory or rely on notes. It allowed us to maximise

2. For an up-to-date thorough review on different tools for automated visual content Analysis for film studies, see (Pustu-Iren et al. 2020).

our time and spot images — and therefore identify patterns — that the human eye could miss. Although far from perfect, Snoop provided us with raw data that we would have not been able to collect manually, at least not with the efficiency and simplicity provided by this AI.

In the *Arclight Guidebook to Media History and Digital Humanities* Hoyt, Hughes, and Acland broadly ponder: “[a]re we innovating and adapting digital tools to address our research questions? Or are we adapting our research interests to fit the available datasets and tools?” (Hoyt et al. 2016: 1). In our case, we did a little of both. As part of our research questions, we deliberated about content overlaps between the films in our corpus and understood that Snoop could help us perform this search. After evaluating the advantages and disadvantages of running the experiments in the conditions that we could achieve now, we decided that it was worth attempting, being aware that our results were going to be approximate and would provide mostly hints of possibilities rather than conclusions.

3 Our Corpus

Between 2017 and 2019, the project “TRANSARCHIVES. Film Heritage and Archival Practices: Past and Present Transcontinental Encounters”, was conducted in the context of a Marie Skłodowska-Curie Fellowship at Sorbonne Nouvelle University.³ With the objective of analysing the circulation of films between the French Cinematheque and those of Argentina, Brazil and Uruguay in the aftermath of the Second World War, this project started constructing a database, based on the correspondence found in the archives of these four institutions and at the FIAF. So far, three hundred and eighty-three films — including fictions and non-fictions of diverse length: feature, medium and short — make up this database including films mentioned in the letters written by Henri Langlois from France; Paulo Emilio Sales Gomes from Brazil; Rolando Fustiñana from Argentina; and Danilo Trelles, Walter Dassori and Eugenio Hintz from Uruguay, between October 1947 and January 1955. Information on this corpus regarding years of production, origins and quantities can be seen in figure 1.

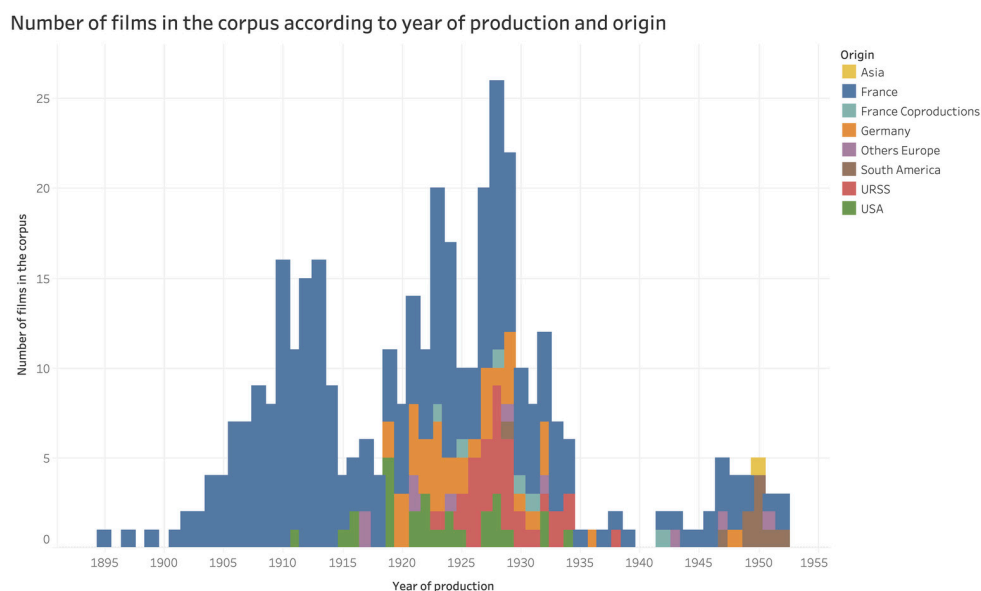


Figure 1: TRANSARCHIVES corpus according to year of production and origin

The mention of a film in these letters does not mean that the film has circulated between these institutions. There are several reasons for which films were mentioned in the correspondence. Among them, there was the

3. TRANSARCHIVES has received funding from the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 746257.

interest to look for missing films, to share a list which could be offered to colleagues, to request specific titles and, indeed, to announce that certain copies were on their way to a given cinemathèque. Nonetheless, the fact that a film was mentioned by this pioneer generation of film archivists suggests that it was part, at least, of an “imaginary collection” of films that deserves attention.⁴

Analysing any corpus using AI first requires access to good quality digital files. This, as we know, is far from being easy. Archives are not necessarily equipped with digital files for most of their films (Bruzzo and Blot-Wellens 2020: 91). Additionally, unless the film archive has its own AI, treating digital collections through this kind of computational system requires sharing the file with the software development team, which copyright restrictions might not always allow (Bruzzo and Blot-Wellens 2020: 93–96; Christensen 2020: 115; Mazzanti 2020: 16). This challenge obliged us to decide whether to wait to reach better experimental conditions or to accept working with *imperfect copies*. The use of the adjective “imperfect”, far from being derogatory or negative, acknowledges that sometimes it is better to move forwards with what we have. Bearing in mind the theoretical reflections of Julio García Espinosa and his advocacy for an “imperfect cinema” in the context of the development of Cuban (García Espinosa 1969), and more broadly, Latin American cinema at the end of the 1960s, we built an appreciation for the possibilities provided by other copy sources.⁵

For over twenty years, the INA have conducted research on different aspects of AI with the aim of simplifying cataloguing processes and helping image retrieval in their collections. In this context, Snoop has been in development since 2014. In addition, as previously mentioned, thanks to their continuous recording, INA keeps copies of films broadcast on French television. These copies can be accessed for research purposes. Our corpus, therefore, would include canonical films which had already undergone a process of *patrimonialisation* (Gauthier 1999: 224–34, Louis 2020, Olesen 2019: 50–66) over several decades. Both the archival and televisual communities have validated these filmic resources, and several film historians have discussed them at length. However, they have not necessarily been through a comparative analysis. This was the main objective of our experiments: to assess whether the AI could find patterns that would allow for comparative analysis to be carried out. Consequently, we had an opportunity to trace shared motifs and images that, in follow-up projects, would allow us to reflect more broadly on how this selection of films has promoted the circulation of certain tropes both in France and abroad.

As a first step, a team of INA documentalists cross-referenced their database with that of TRANSARCHIVES. The search found that one hundred and forty-four films (38%) have been aired on television,⁶ but of this amount only sixty-four films (17% of the total) were stored in digital files that Snoop could readily treat. Although the corpus was smaller than we had anticipated, it was still big enough to run our experiments.

Our corpus ended up being randomly chosen, based on the availability of the digital files and not following parameters of year of production, origin, genre or authorship. Rather than trying to create a more homogeneous or a larger corpus, we considered that this random sample would allow us to try the AI “in the wild”, and therefore emphasise the possibilities offered by this tool. Studying cinema through a television archive posed a few challenges. First, that the files might include advertising or other moving images that do not belong to the film and needed to be discarded from the findings. Second, that data on the version of the film is not always available, an issue that we are aware is of concern in the field (Cherchi Usai 2012: 535, Flueckiger et al. 2016, Heftberger 2012, 2018: 4). Third, that it is not possible to know the criteria applied for adjusting the aspect-ratio of the film to fit that of the television. Another challenge, as mentioned above, was that our corpus was made of archival films and therefore Snoop could not perform optimally where the quality of the images did not allow.

-
4. The idea of “imaginary collection” establishes a connection with the “imaginary museum” of André Malraux, which intended to be as broad as possible (Malraux 1947). For a comprehensive study of Malraux’s concept, see (Charlotte Wasser 2019). This notion has also been widely linked to Henri Langlois. To commemorate the centenary of his birth the French Cinemathèque organised an exhibition called “Le musée imaginaire d’Henri Langlois.” See Païni 2014.
 5. Based on an analysis by Michael Chanan, Catherine Grant brings the idea of “imperfect digital humanities” to bear on amateur and lo-fi digital manifestations which choose to break with the commodification of corporate-funded technological research and commercial academic publishing, enabling mere ‘spectators’ to become ‘agents’ (Grant 2012).
 6. INA has a team of permanent staff members who help researchers. In this case, the team did thorough research to compare the list from TRANSARCHIVES (considering differences in titles’ languages, dates and directors’ names) to provide the definite list of films that had been broadcast on French television.

Exhibition of corpus on French TV

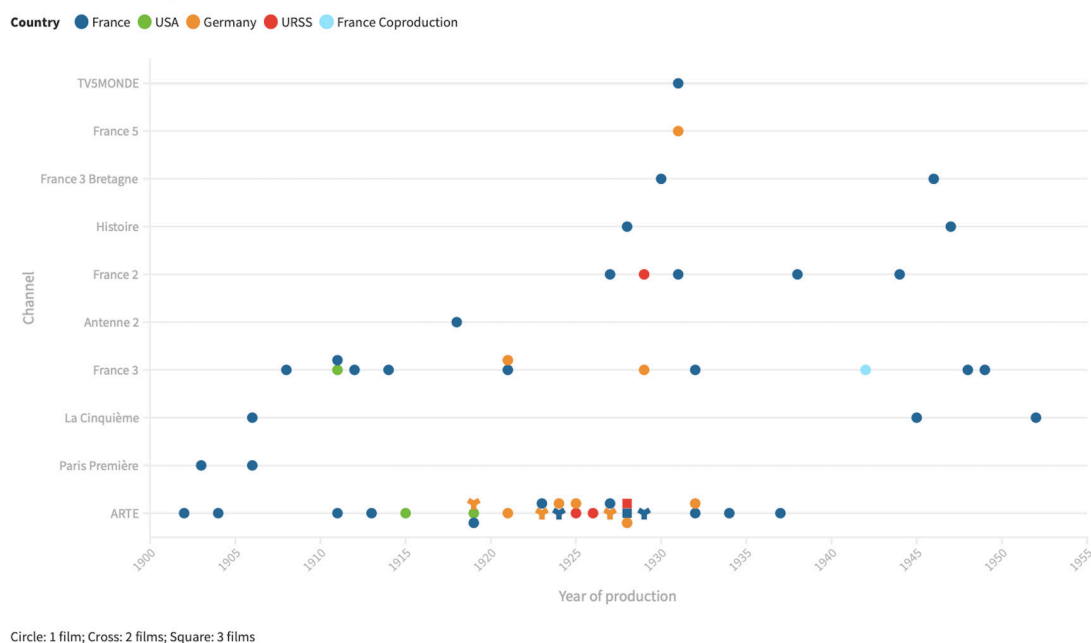


Figure 2: Visualisation on our corpus, including the years of production, origins and French television channels that broadcast the films

4 Snoop and the Methodology Behind our Experiments

Snoop allows users to search for any given visual content in large video collections, and then build models for the content that users aim to identify. It is a large-scale content-based image and video search engine whose main features are:⁷

- the extraction and efficient indexing of visual features from pictures or videos. Visual features are mathematical representations: vectors. These vectors are extracted by specific algorithms. There are two types of algorithms, for instance:
 - hand-crafted algorithms, which are designed by humans: Sift (Lowe 1999) or Gist (Oliva and Torralba 2001),
 - or learned through deep learning methods: InceptionV2 (Szegedy et al. 2015) or ResNet (He et al. 2015);
- the identification of similar images or keyframes through approximate k-nearest neighbours (Joly and Buisson 2008);
- the supervised recognition of trained visual concepts; and
- relevant feedback modules to create datasets for specific classes or visual concepts.

The supervised recognition of trained visual concepts has been the main tendency in academic writing and commercial products during the last decade due to its high efficiency. This efficiency is based on the Machine Learning algorithms that are trained with thousands of examples for each concept. To create a collection composed of hundreds of thousands of concepts, we would need to annotate millions of images. This is an operation that requires expensive human resources. Moreover, this functionality is very useful if the user

7. Snoop is notably the visual search engine used by the Pl@ntNet (<https://plantnet.org/>) application and in this context, it has a very large audience made up of more than six million users.

already knows what they want to retrieve in their document collection, such as people, faces, dogs and cars. This classical functionality can be retrieved in different toolkits like those aforementioned. As soon as the user has annotated these sample collections, they launch for a few days the AI algorithms to learn the concept models. The main problem with supervised recognition occurs when the user wants to update the collection of visual concepts. Any such update involves the annotation of thousands of new examples for each new concept followed by a re-launch of the estimation process. Thus, as soon as the user discovers a new set of visual concepts, they have to spend a few days training the AI before being able to apply the model.

In contrast, Snoop uses the “active learning” method that permits simpler and more adaptable user interaction (Musik and Zeppelzauer 2018). Snoop therefore relies on relevant feedback strategy, which is based on specific active learning methods. This relevant feedback strategy avoids excessive human resource costs and enhances the discovery of contents for research corpora. The relevant feedback (Li and Allinson 2013, Tzelepi and Tefas 2016) module has been developed in Snoop to help users discover new visual concepts and to create models to automatically retrieve these new concepts. This module is illustrated in figure 3.

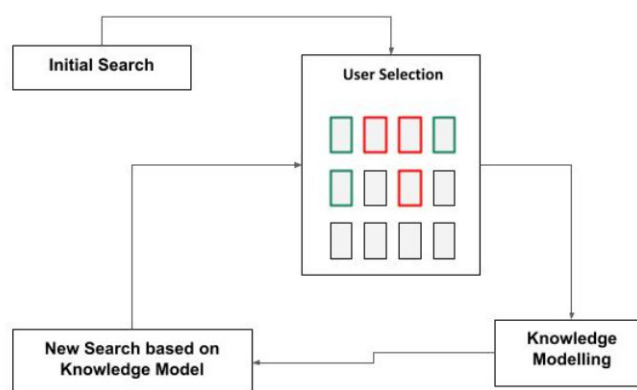


Figure 3: Relevant Feedback module

The user starts the process with an initial search of one image or a set of images. Snoop retrieves the most similar images or video excerpts indexed during the first step (the extraction and efficient indexing of visual features from the corpus). Among these results, the user selects positive and negative samples. The “positive samples” represent what the user wants to retrieve, and the “negative samples” what does not match the search. These positive and negative samples are used by a machine learning algorithm to model the knowledge provided by the user’s annotations. The negative samples are very useful to obtain an accurate estimation of the knowledge model. This knowledge model is used to automatically generate queries in order to retrieve new images or video excerpts. These new images or video excerpts are proposed to the user for another annotation step. During this step, the user again chooses positive and negative samples to transfer knowledge to the machine learning algorithm. This process is repeated until the user is satisfied by the retrieved visual documents and the knowledge model’s stability.

To create the visual concepts of the experiments, the initial search is the most important step of the relevant feedback module. This first step defines the objects intended for retrieval. In our experiments we used two complementary strategies:

- First, we defined a set of visual concepts that we selected to be analysed in our documents. We based our choices on visual tropes that we identified by watching the films, on the book *Images du cinéma français* (Védère 1945), and on motifs found in paintings contemporary to the films (especially from the avant-garde). Among our choices, we included: contemporaneous iconography — most of which was related to modernity — such as the identification of crowds and technological advances such as planes, trains and cars; objects that evoke war and authority (a variety of uniforms, guns); older means of transport (carts, horses); natural elements such as water and images related to it such as swimmers, waves and boats; leisure activities such as dancing; cinematic tropes such as a couple kissing.

- The second set of choices was based on serendipity. In order to discover visual contents, no *a priori* was set and we simply navigated the corpus with a random selection of video excerpts. As soon as we discovered an interesting visual concept, we started the relevant feedback process. This approach was good to explore, for instance, fashion elements, such as hats, embroidered dresses and jackets, and haircuts.

This relevant feedback process allows us to apply the “Human in the Loop” (Monarch 2021) approach which provides rich data to the machine in order to define machine learning models for the visual concepts very accurately.

5 Results and New Lines of Enquiry

As indicated, the main objective of our experiments was to assess whether the AI could find patterns that would allow for comparative analysis to be carried out within our corpus of canonical films. This section, therefore, focuses on highlighting the possibilities offered by AI to film historians. At this stage we cannot present all the classes we have built or provide thorough analyses, though this could be certainly achieved in follow up projects. However, we will briefly develop the study of the class built around images of crowds and demonstrate how an approach involving an AI would differ from previous ones. We will also suggest lines of research that might be opened thanks to the classes that retrieved close ups of hands and images of men in uniforms. The class encompassing crowds was the richest of our experiments. To build this class we presented Snoop with only five photographs taken from *Gallica*, the open access database of the French National library (BnF), and then started choosing positive and negative samples from the images Snoop proposed.⁸ We ended up retrieving crowds from twenty-three films of our corpus (35%), which although not being negligible, is likely to be incomplete.⁹

Although crowds are present in films from the beginning of cinema, film scholars have devoted only modest attention to them. Eric Faden’s study (2001), for instance, emphasises the relationship between the development of technology — early cinema, sound and digital images — and the image of the crowd to explore different “crowd control mechanisms”, based on the work of James R. Beniger (Faden 2001). Methodologically, Faden chooses to support his in-depth study of the relationship between the control of technologies and that of the crowd’s image with a few examples from specific film sequences. Although different in approach, the large timeframe chosen by Faden is relevant to our study since it demonstrates that far from being an exception or an element constrained to a specific historical point in time, the inclusion of crowds in film narratives have been widespread from the very beginning.

There are other approaches to the study of crowds in cinema, for example, those that Lesley Brill (2006) and Georges Didi-Huberman (2016) have proposed. While Brill based his work on that of Elias Canetti, to analyse “filmic representations of crowd qualities and dynamics and representations of the crowd’s usually evil twin, power” (Brill 2006, 4), through analyses of specific films, Georges Didi-Huberman, focuses on how masses reveal collective emotions in *Battleship Potemkin* (Serguei M. Eisenstein, 1925) (Didi-Huberman 2016: 195–202). In both cases, individual in-depth analyses are privileged. Although this approach could certainly complement a computer-aided study of a larger corpus, rather than exploring the behaviour, construction or relevance of the crowd within the narrative of an individual film, new tools trigger different questions and offer the possibility of establishing comparative approaches within a larger number of films. Indeed, this is clearly seen in Eva Hielscher’s study that analysed a corpus of sixty-seven City Symphonies to compare the main characteristics of these “experimental–documentary city films of the 1920s and 1930s” (Hielscher 2020: 1).

For her work, Hielscher used, among others computer tools, the annotation system ELAN to manually keep

8. <https://gallica.bnf.fr/>.

9. We know, for instance, that Snoop did miss some important films in our corpus, such as *Grand Illusion* (*La Grande Illusion*, 1937), in which crowds have already been analysed as being part of “critical narrative points” (Brill 2006: 9). However, as previously discussed, we focus here on the findings from our first round of experiments, rather than trying to find explanations for Snoop missing images of crowds. Refining its algorithm to miss fewer occurrences has not been part of our aims within the timeframe of our experiments.

track of several characteristics of four canonical city symphonies.¹⁰ She then compared some shared features with the rest of the corpus, which were explored through ordinary viewings (Hielscher 2020: 13). Hielscher's methodology allowed her to spot shared motifs, and the crowd was among these. Her research concludes that in forty-four of the sixty-seven titles, "there are not too many 'real' crowd shots with masses of people. Instead, we can often speak of an accumulation technique, by with (sic) the impression of crowds is created visually through editing" (Hielscher 2020: 14). Although the size of Hielscher's corpus is almost the same as ours (ours has three fewer films but includes feature length films) and we are also looking for shared motifs, both the tools and the selection of the corpus are different. While Hielscher's corpus is made of films described under the label "city symphony", the films in our corpus share no thematic consistency. They share the fact of having been mentioned by film archivists in letters written at a specific time. In addition, we have not selected subgroups of films that Snoop could use to build a strong model that would later be applied to the rest. All of the films were on an equal footing and all of them served equally to build the model. Using an AI implies changing the methodology and shifting to new ways of studying film history.

In the case of retrieving crowds, our objective was not to analyse the crowd's role in the narrative of a specific film, or whether and how an *a priori* constructed group of films share techniques, styles, objectives, etcetera, when depicting crowds. Introducing an AI in a study like the one we are proposing, allowed for finding similarities — and dissimilarities — in films that would not have necessarily been studied together without engaging in hundreds of hours of film watching or laborious annotations, which in practice can only be performed in a restricted number of films.

We were thus able to retrieve similar frames in disparate films all of which show how the spectator's attention is directed towards a specific point using images of crowds. A first set of images (figure 4) shows crowds escaping, running towards something or simply with their backs turned towards the spectator so that their attention goes beyond the crowd.

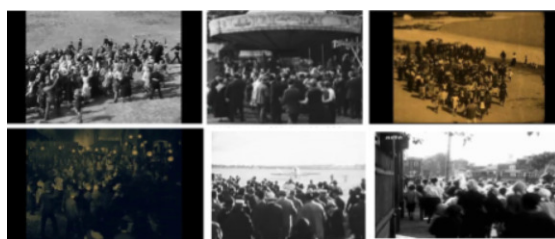


Figure 4: From left to right and top to bottom: *The Woman who Dared* (*Le ciel est à vous*, Jean Grémillon 1944), *The Red Head* (*Poil de carotte*, Julien Duvivier 1932), *I accuse* (*J'accuse*, Abel Gance 1919), *The late Mathias Pascal* (*Feu Mathias Pascal*, 1924), *Money* (*L'Argent*, Marcel L'Herbier 1928), *Strike* (*Stachka*, Sergueï Eisenstein 1925)

The second set (figure 5) shows images that were shot from a much closer distance and therefore allow us to read the faces of individual people who together express feelings of celebration, astonishment or disapproval. In this set, the point of view is directed to the off-screen space which could have been introduced before the crowd or presented later, depending on whether the spectator is expected to confirm or anticipate these reactions.

The third set (figure 6) includes in the same frame the crowd and the element that attracts its attention. This was the only way of showing a crowd used in three of the earliest films of our collection (two by Méliès and one by Capellani). However, as we have already pointed out, at this stage we know that plenty of examples escaped Snoop's attention so we would not feel comfortable providing conclusive observations. Rather, we prefer to provide examples for possible further studies. In-depth analyses of these findings could shed light on the development of cinematic language that have been taken for granted or only explored through consistent and controlled cases.

10. For a comprehensive review of ELAN and other annotation tools, see (Flueckiger 2017, Melgar Estrada et al. 2017, Pustu-Iren et al. 2020).



Figure 5: From left to right and top to bottom: *Metropolis* (Fritz Lang 1927), *Max Bullfighter* (*Max toréador*, Max Linder 1913), *Money*, *Napoleon* (Abel Gance 1927), *Boys' School* (*Les disparus de St Agil*, Christian-Jaque 1938)



Figure 6: From left to right and top to bottom: *Arsenal* (Alexandre Dovjenko 1928), *Napoleon*, *The Weavers* (*Die Weber* William Dieterle, Friedrich Zelnik, 1927), *The Impossible Voyage* (*Voyage à travers l'impossible*, Georges Méliès 1904), *A Trip to the Moon* (*Le voyage dans la lune* Georges Méliès 1902), *Aladdin* (Albert Capellani 1906)

As mentioned above, we constructed one class around close-ups of hands and although Snoop retrieved few cases that we marked as positive, we were impressed by the aesthetic beauty of the examples of these expressive hands from films of the late 1920s and early 1930s (figure 7). We anticipate that rich lines of aesthetic — and political — analyses could be developed by comparing these images with contemporary art works such as *The Hands of Antonin Artaud* (Man Ray circa 1925), *Two Crossed Hands* (Pablo Picasso 1921), *The Hand Has Five Fingers* (John Heartfield 1928) and *The Hand* (Salvador Dali 1930), to name only a few.



Figure 7: From left to right and top to bottom: *Diary of a Lost Girl* (*Tagebuch einer Verlorenen*, Georg Wilhelm Pabst 1929), *M* (Fritz Lang 1931), *The Love of Jeanne Ney* (*Die Liebe der Jeanne Ney*, Georg Wilhelm Pabst 1928), *The Wheel* (*La Roue*, Abel Gance 1923), *iFinis Terrae* (Jean Epstein 1929), *October* (Sergueï Eisenstein 1928), *The Blue Express* (*Goluboj ekspress*, Ilya Trauberg 1928), *Great Illusion* (*La grande illusion* Jean Renoir 1937), *Arsenal*

Likewise, we envisage that scholars interested in representations of masculinity, could find the results obtained around uniforms worn by men (from the army, policemen, security guards) relevant. Showing that a non-negligible number of men have worn military uniforms throughout the history of cinema is not a minor detail when thinking about the construction of masculine imaginaries. Colleagues in this research area could benefit from these AI findings that provide a different or complementary perspective (figure 8).



Figure 8: From left to right and top to bottom: *October*, *Battleship Potemkin* (*Bronenosets Potyomkin* Serguei Eisenstein 1926), *The Blue Express*, *The New Men* (*Les nouveaux messieurs* Jacques Feyder 1929), *Man about Town* (*Le silence est d'or*, René Clair 1927), *Freedom for Us* (*À nous la liberté*, René Clair 1931), *Napoleon*, *The Little Match Girl* (*La petite marchande d'allumettes*, Jean Renoir 1928), *The Dance of Death* (*La danse de mort*, Marcel Cravenne 1948)

In further studies, all these searches could be refined, for instance, loading only the films which retrieved similarities in a first search and running Snoop again. Proceeding this way, the search would be more controlled and the possibilities of analysis richer. Loading specific images from the corpus would also help Snoop carry out more precise searches and retrieve other relevant examples. Although approximate, our findings demonstrate the possibilities that AI could offer to film historians in tracing stylistic conventions, providing precise examples and engaging in comparative analysis that theoretically might not have been thought of as worth pursuing.

6 Conclusion

This article aimed to explore the possibilities offered by AI for film historians to develop new methods and research questions. Although having been developed and used for a few years now, film historians have not greatly explored AI. Our study shows that certain pre-conditions (which are not necessarily easy to obtain) are necessary to engage AI in such research. Finding an archive that has developed an efficient AI and that also has digital film copies accessible could be challenging. In addition, concessions might need to be made in terms of both quality and quantity. In our case, we were aware that our copies were “imperfect copies” and that our corpus was smaller than we would have liked. However, we decided that it was worth trying a first round of experiments to have an idea of AI’s possible uses.

In practical terms, although the fantasy of AI doing the work for us might exist, there is laborious hands-on selection involved throughout the whole process which led us to become thoroughly acquainted with the corpus. In addition, our method could be used as a complement to those that we have been using for decades. An AI could run a primary search to start reducing a corpus, followed by an in-depth individual analyses, therefore oscillating between distant and close viewing.¹¹ We could also follow Hielscher’s methodology and create a themed or genre-based corpus and build a model using a discrete core of canonical cases (Hielscher 2020). In fact, by carrying out the experiments, we realised the myriad of possibilities offered by this technology and the several methods that could be followed to use it. The fact that on this occasion we focused on qualitative rather than quantitative analyses does not indicate that AI cannot be used in the converse manner or for both at the same time. The choice of analysis should be related to the possibilities offered by the tool as well as by

11. For an article that reflects on qualitative computational analysis of audio-visual media, and the advantages and shortcomings of distant and close reading, even if using other tools than AI, see (Bakels et al. 2020).

the specific research aims of each project. Now is the time to continue refining the tool and exploring further hypotheses. Close collaborations between computing scientists, film scholars and archivists are required to carry out this kind of project, and in this respect our experiments proved very successful, with close cooperation and smooth communication processes. We understand that AI is here to stay and that the more we use it, the more we will be able to identify its benefits in developing our research activities and to continue refining our questions and methods.

7 Bibliography

- Arnold, Taylor, and Lauren Tilton (2019). "Distant Viewing: Analyzing Large Visual Corpora". *Digital Scholarship in the Humanities* 34 (Supplement_1): 13–16 <https://doi.org/10.1093/llc/fqz013>.
- Arnold, Taylor, Lauren Tilton, and Annie Berke (2019). "Visual Style in Two Network Era Sitcoms". *Journal of Cultural Analytics* <https://doi.org/10.22148/16.043>.
- Bakels, Jan-Hendrik, Matthias Grotkopp, Thomas Scherer, and Jasper Stratil (2020). "Matching Computational Analysis and Human Experience: Performative Arts and the Digital Humanities". *DHQ: Digital Humanities Quarterly* 14(4) <http://www.digitalhumanities.org/dhq/vol/14/4/000496/000496.html> (last accessed 03-03-21).
- Bardèche, Maurice, and Robert Brasillach (1935). *Histoire Du Cinéma*. Paris: Denoël et Steele.
- Bermeitinger, Bernhard, Sebastian Gassner, Siegfried Handschuh, Gernot Howanitz, Rik Radisch, Malte Rehbein (2019). "Deep Watching: Towards New Methods of Analyzing Visual Media in Cultural Studies". *Data-verseNL* <https://doi.org/10.34894/HOSP8P>.
- Brill, Lesley (2006). *Crowds, Power, and Transformation in Cinema*. Detroit: Wayne State University Press.
- Bruzzo, Mariona, and Camille Blot-Wellens (2020). "A Curatorial Approach to Making Cinematic Heritage Available Online". In *I-Media-Cities: Innovative e-Environment for Research on Cities and the Media*, edited by Teresa M Sala and Mariona Bruzzo, 89–100. Barcelona: Edicions de la Universitat de Barcelona.
- Burghardt, Manuel, Adelheid Heftberger, Johannes Pause, Niels-Oliver Walkowski, and Matthias Zeppelzauer (2020). "Film and Video Analysis in the Digital Humanities – An Interdisciplinary Dialog". *Digital Humanities Quarterly* 14(4) <http://www.digitalhumanities.org/dhq/vol/14/4/000532/000532.html> (last accessed 01-03-21).
- Charlotte Wasser (2019) "Le Musée Imaginaire d'André Malraux: un anti-musée? Une nouvelle forme de musée par l'édition". *Présence d'André Malraux sur la Toile* <https://malraux.org/13509-2/> (last accessed 30-05-21).
- Cherchi Usai, Paolo (2012). "Early Films in the Age of Content; or, "Cinema of Attractions" Pursued by Digital Means". In *A Companion to Early Cinema*, edited by André Gaudreault, Nicolas Dulac, and Santiago Hidalgo, 527–49. Malden: Wiley-Blackwell.
- Christensen, Thomas C. (2020). "RE-USE. How Might We Unlock and (Re-)Use What the European Film Archives Hold?" In *I-Media-Cities: Innovative e-Environment for Research on Cities and the Media*, edited by Teresa-M Sala and Mariona Bruzzo, 115-118. Barcelona: Edicions de la Universitat de Barcelona.
- Didi-Huberman, Georges (2016). *Peuples En Larmes, Peuples En Armes*. Paris: Les Éditions de Minuit.
- Elsaesser, Thomas (2012). "Is Nothing New? Turn-of-the-Century Epistemes in Film History". In *A Companion to Early Cinema*, edited by André Gaudreault, Nicolas Dulac, and Santiago Hidalgo, 587–609. Malden: Wiley-Blackwell.
- Faden, Eric S. (2001). "Crowd Control: Early Cinema, Sound, and Digital Images". *Journal of Film and Video* 53(2/3): 93–106 <https://www.jstor.org/stable/20688359> (last accessed 1-10-21).
- Flueckiger, Barbara (2017). "A Digital Humanities Approach to Film Colors". *The Moving Image: The Journal of the Association of Moving Image Archivists* 17(2): 71 <https://doi.org/10.5749/movingimage.17.2.0071>.
- Flueckiger, Barbara, Franziska Heller, Claudy Op den Kamp, and David Pfluger (2016). "'Digital Desmet': Translating Early Applied Colors". *The Moving Image: The Journal of the Association of Moving Image Archivists* 16(1): 106 <https://doi.org/10.5749/movingimage.16.1.0106>.
- García Espinosa, Julio (1969 [1971]). "Por Un Cine Imperfecto". *Revista Cine Cubano* (66/67).
- Gaudreault, André (2006). "From 'Primitive Cinema' to 'Kine-Attractography'". In *The Cinema of Attractions Reloaded*, edited by Wanda Strauven, 85–104. Amsterdam: Amsterdam University Press.

Gaudreault, André, Nicolas Dulac, and Santiago Hidalgo (2012). "Introduction". In *A Companion to Early Cinema*, edited by André Gaudreault, Nicolas Dulac, and Santiago Hidalgo, 1–11. Malden: Wiley-Blackwell.

Gaudreault, André, and Tom Gunning (1989). "Le cinéma des premiers temps, un défi à l'histoire du cinéma?" In *L'histoire du cinéma: nouvelles approches*, edited by Jacques Aumont, André Gaudreault, and Michel Marie, 49–63. Paris: Publications de la Sorbonne.

Gauthier, Christophe (1999). *La Passion Du Cinéma: Cinéphiles, Ciné-Clubs et Salles Spécialisées à Paris de 1920 à 1929*. Paris: Ecole nationale des chartes: Association française de recherche sur l'histoire du cinéma.

Grant, Catherine (2012). "Film and Moving Image Studies: Re-Born Digital? Some Participant Observations". *Frames Cinema Journal* (1) <http://framescinemajournal.com/article/re-born-digital/> (last accessed 04-04-21).

Gunning, Tom (2006). "Attractions: How They Came into the World". In *The Cinema of Attractions Reloaded*, edited by Wanda Strauven, 31–39. Amsterdam: Amsterdam University Press.

He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun (2015). "Deep Residual Learning for Image Recognition". *arXiv:1512.03385 [cs]* <http://arxiv.org/abs/1512.03385> (last accessed 18-10-21).

Heftberger, Adelheid (2012). "Ask Not What Your Web Can Do For You – Ask What You Can Do For Your Web! Some Speculations about Film Studies in the Age of the Digital Humanities". *Frames Cinema Journal* (1) <https://framescinemajournal.com/article/ask-not-what-your-web-can-do-for-you/> (last accessed 22-04-21).

— — — (2018). *Digital Humanities and Film Studies: Visualising Dziga Vertov's Work*. Cham: Springer International Publishing.

Hielscher, Eva (2020). "The Phenomenon of Interwar City Symphonies: A Combined Methodology of Digital Tools and Traditional Film Analysis Methods to Study Visual Motifs and Structural Patterns of Experimental-Documentary City Films". *Digital Humanities Quarterly* 14(4) <http://www.digitalhumanities.org/dhq/vol/14/4/000495/000495.html> (last accessed 05-10-21).

Hoyt, Eric, Kit Hughes, and Charles R. Acland (2016). "A Guide to the Arclight Guidebook". In *The Arclight Guidebook to Media History and the Digital Humanities*, edited by Charles R. Acland and Eric Hoyt, 1–29. Sussex: Reframe Books <http://projectarclight.org/book/> (last accessed 1-10-21).

Joly, Alexis, and Olivier Buisson (2008). "A Posteriori Multi-Probe Locality Sensitive Hashing". In *Proceeding of the 16th ACM International Conference on Multimedia – MM '08*, Vancouver, British Columbia, Canada: ACM Press <http://doi.org/10.1145/1459359.1459388>.

Kaplan, Andreas, and Michael Haenlein (2019). "Siri, Siri, in My Hand: Who's the Fairest in the Land? On the Interpretations, Illustrations, and Implications of Artificial Intelligence". *Business Horizons* 62(1): 15–25. <https://doi.org/10.1016/j.bushor.2018.08.004>.

Latsis, Dimitrios, and Grazia Ingravalle (2017). "Guest Editors' Foreword: Digital Humanities and/in Film Archives". *The Moving Image: The Journal of the Association of Moving Image Archivists* 17(2): xi <https://doi.org/10.5749/movingimage.17.2.00xi>.

Li, Jing, and Nigel M. Allinson (2013). "Relevance Feedback in Content-Based Image Retrieval: A Survey". In *Handbook on Neural Information Processing*, Intelligent Systems Reference Library, edited by Monica Bianchini, Marco Maggini, and Lakhmi C. Jain. Berlin, Heidelberg: Springer Berlin Heidelberg, 433–69 <http://link.springer.com/10.1007/978-3-642-36657-413>. (last accessed 11-10-21).

Louis, Stéphanie E. (2020). *La Cinémathèque-Musée: Une Innovation Cinéphile Au Coeur de La Patrimonialisation Du Cinéma En France (1944-1968)*. Paris: AFRHC.

Lowe, D.G. (1999). "Object Recognition from Local Scale-Invariant Features". In *Proceedings of the Seventh IEEE International Conference on Computer Vision*, Kerkyra, Greece: IEEE, 1150–57 vol.2. <http://doi.org/10.1109/ICCV.1999.790410>.

Malraux, André (1947). *Psychologie de l'art. Le Musée Imaginaire*. Paris: Albert Skira Éditeur.

- Mazzanti, Nicola (2020). "SPREAD. (RE)SEARCH". In *I-Media-Cities: Innovative e-Environment for Research on Cities and the Media*, edited by Teresa-M Sala and Mariona Bruzzo, 15–20. Barcelona: Edicions de la Universitat de Barcelona.
- Melgar Estrada, Liliana, Eva Hielscher, Marijn Koolen, Christian G. Olesen, Julia Noordegraaf, Jaap Blom (2017). "Film Analysis as Annotation: Exploring Current Tools". *The Moving Image: The Journal of the Association of Moving Image Archivists* 17(2): 40–70 <https://doi.org/10.5749/movingimage.17.2.0040>.
- Mitry, Jean (1968). *Histoire du cinéma 1. 1895-1915*. Paris: Ed. universitaires.
- Monarch, Robert (2021). *Human-in-the-Loop Machine Learning: Active Learning and Annotation for Human-Centered AI*. Sherlter Island, NY: Manning Publications.
- Mulvey, Laura (2006). *Death 24x a Second: Stillness and the Moving Image*. London: Reaktion Books.
- Musik, Christoph, and Matthias Zeppelzauer (2018). "Computer Vision and the Digital Humanities: Adapting Image Processing Algorithms and Ground Truth through Active Learning". *VIEW Journal of European Television History and Culture* 7(14): 1–14 <https://doi.org/10.18146/2213-0969.2018.jethc153>.
- Olesen, Christian (2019). "'This Is Our First Big Experiment': Paris 1900 (1947) and the Eye Filmmuseum's Early Collection-Building". *Early Popular Visual Culture* 17(2): 207–17 <https://doi.org/10.1080/17460654.2019.1697113>.
- Oliva, Aude, and Antonio Torralba (2001). "Modeling the Shape of the Scene: A Holistic Representation of the Spatial Envelope". *International Journal of Computer Vision* 42(3): 145–75 <http://doi.org/10.1023/A:1011139631724>.
- Païni, Dominique (2014). *Le Musée Imaginaire d'Henri Langlois*. Paris: Flammarion: Cinémathèque française.
- Pustu-Iren, Kader, Julian Sittel, Roman Mauer, Oksana Bulgakowa, Ralph Ewerth (2020). "Automated Visual Content Analysis for Film Studies: Current Status and Challenges". *Digital Humanities Quarterly* 14(4) <http://www.digitalhumanities.org/dhq/vol/14/4/000518/000518.html> (last accessed 12-09-21).
- Rombes, Nicholas (2017). *Cinema in the Digital Age*. Revised edition. London, New York: Wallflower Press.
- Sadoul, Georges (1946). *Histoire Générale Du Cinéma. Tome 1: L'Invention Du Cinéma, 1832-1897*. Paris: Éditions Denoel.
- (1947). *Histoire Générale Du Cinéma. Tome 2: Les Pionniers Du Cinéma, 1897-1909*. Paris: Éditions Denoel.
- (1948). *Le Cinéma. Son Art, Sa Technique, Son Économie*. Paris: Éditions Denoel.
- (1949). *Histoire Du Cinéma Mondial*. Paris: Flammarion.
- Szegedy, Christian, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, Zbigniew Wojna (2015). "Rethinking the Inception Architecture for Computer Vision". *arXiv:1512.00567 [cs]* <http://arxiv.org/abs/1512.00567> (last accessed 10-10-21).
- Tadeo Fuica, Beatriz (2019). "Tracing Past Exchanges between European and South American Cinematheques: A Key to Understanding the Impact of Sharing". *Illuminace. The journal of film theory, history, and aesthetics* 31(1): 27–43.
- Tzelepi, Maria, and Anastasios Tefas (2016). "Relevance Feedback in Deep Convolutional Neural Networks for Content Based Image Retrieval". In *Proceedings of the 9th Hellenic Conference on Artificial Intelligence*, Thessaloniki Greece: ACM, 1–7 <https://doi.org/10.1145/2903220.2903240>.
- Védres, Nicole (1945). *Images Du Cinéma Français*. Paris: Les éditions du chêne.

Beatriz Tadeo Fuica – Université Sorbonne Nouvelle, Paris 3 (France)

✉ beatriz.tadeo-fuica@sorbonne-nouvelle.fr

She is a research associate at the IRCAV, Université Sorbonne Nouvelle – Paris 3, where she developed the Marie Skłodowska-Curie project “TRANSARCHIVES: Film Heritage and Archival Practices: Past and Present Transcontinental Encounters”. She holds a PhD from the University of St Andrews (UK).

Olivier Buisson – French Audiovisual National institute (INA) (France)

He is a senior researcher at the INA’s Research Department. For 20 years, he has been leading research activities in Artificial Intelligence to explore and annotate very large corpora of images and videos. His specific research topics are: Interactive Machine Learning, object or content retrieval, discovery and mining, large scale similarity search (High-dimensional indexing structures, Hashing methods). He holds a PhD from Université de La Rochelle (France).

Claude Mussou – French Audiovisual National institute (INA) (France)

She is the head of INA-thèque, INA’s department in charge of making the collections available for research purposes and developing academic usage. She graduated from a 2-year training program in Digital Humanities at Sciences Po in Paris (France).