Journal of Scandinavian Cinema
Volume 12 Number 3
© 2022 Intellect Ltd Feature Article. English language.
https://doi.org/10.1386/jsca_00075_1
Received 3 November 2021; Accepted 16 February 2022

feature article

**MARIA ERIKSSON**
Umeå University and Basel University

**TOMAS SKOTARE**
Umeå University

**PELLE SNICKARS**
Lund University

# Understanding Gardar Sahlberg with neural nets: On algorithmic reuse of the Swedish SF archive

**ABSTRACT**

*In this article, we re-trace the history of the Swedish SF archive and reflect on how this collection of historic newsreels has been reappropriated and remixed throughout more recent media history. In particular, we focus on the work of director and film historian Gardar Sahlberg, who made extensive use of the SF archive, first in a series of documentary films, then in a number of historical TV programmes. We are interested in how historic film footage travels and circulates through time, but foremost we explore how algorithms can help identify instances of audio-visual reuse in large datasets. Hence the article discusses algorithmic ways of examining archival film reuse, introducing a method for mapping video reuse with the help of*

**KEYWORDS**

film archives
archival reuse
Video Reuse Detector
AI
convolutional neural nets
computational film studies

*artificial intelligence or more precisely machine learning that uses so-called convolutional neural nets. The article presents the Video Reuse Detector (VRD), a tool that uses machine learning to identify visual similarities within a given audiovisual database such as the SF archive.*

In early January 1964, the CEO of Sveriges Radio (SR) (Radio Sweden) was pleased to announce that his organization had finally acquired the old newsreel archive from the production company Svensk Filmindustri (SF). One million metres of film dating from 1897 to 1960 had been purchased for 3 million Swedish crowns. At the time Swedish television was still a branch of SR, and the idea of acquiring a film archive was to use old footage in new TV programmes. The deal was explosive, not only because of the number of preserved nitrate prints; the SF archive, or SF's journalfilmarkiv (SF's newsreel archive) as it was originally known, would now find a novel audience in another medium: television. Extensive reuse of the archive promised an array of entertaining and informative TV programmes. Transferring this remarkable film archive to a public broadcaster was of 'significant public interest', at least according to Kenne Fant, then CEO at SF. The SF archive would give 'invaluable access' to a mediated past for historical TV programmes, and the archive promised to become a 'film mine to the delight of [public service] license payers' (Anon. 1964: n.pag.).

Researchers on Swedish film history have previously provided a detailed description of this major film archival settlement between a national film industry in decline and a public service broadcaster on the rise (Snickars 2006; Vesterlund 2012, 2019). The arrival of television had slashed cinema attendance in half and the film industry was hardly in favour of films being screened on TV, but with old newsreels an exception was made. Arranging the contractual agreement between SF and SR, however, became complicated. During 1963 the Swedish national archivist Ingvar Andersson had tried hard to estimate the value of the SF archive, working together with Harry Schein, CEO of the Swedish Film Institute (SFI). Since old film material was rather popular, calculations became cumbersome. SF did make money on sales of screening rights; footage for American television, for example, was sold for SEK 2754 per minute (Fant 1963). The major problem was the state of the archive; a substantial part of it contained nitrate negatives. Hence, the film material needed to be moved to and restored at SR, a prerequisite before any actual reuse and programme production could take place. Nevertheless, the main principle behind the agreement was archival reuse – in both theory and practice.

In this article, we re-trace the history of the SF archive and reflect on how this particular collection of historic newsreels, short films and early film footage has been reappropriated and remixed throughout more recent media history. In particular, we focus on the work of the writer, director and film historian Gardar Sahlberg, who made extensive use of the SF archive, first in a series of documentary films, then in a number of historical TV programmes. We are interested in the way historic film footage travels and circulates through time, but foremost we want to explore how algorithms can help identify instances of audio-visual reuse in large datasets. Given that it would be a tedious task for a film historian to detect all forms of reuse of the SF archive in Sahlberg's film and TV productions, we experiment with deploying a set of archival algorithms to help us identify reappearing shots and sequences – or rather video frames in our terminology – in Sahlberg's production. This is possible since the complete SF archive was digitized in the early 2000s in a collaboration between Sveriges Television (SVT) (Swedish Television) and the Swedish National Archive of Recorded Sound and Moving Images (SLBA) (Snickars 2003: n.pag.). While algorithms make it possible to detect every single frame from the SF archive that Sahlberg reused – at least in theory – we focus on one of his compilation films as a case study, *Brefven från Stockholm: En film om sommaren 1909* ('Letters from Stockholm: A film about the summer of 1909') (Sahlberg 1964) and computationally examine how a minor part of the SF archive – footage from 1909, to be exact – reappears in this film.

The purpose of the article is threefold. First, we unpack the early history and afterlife of the SF newsreel archive through Sahlberg's work. Reuse is the notion that our discussion revolves around, a term similar to archival practices focusing on how 'remediation and recontextualization of footage contribute to the circulation of audiovisual memory' (Brunow 2017: 98). Second, we discuss algorithmic methods of examining archival film reuse, and by doing so introduce a method for mapping video reuse with the help of artificial intelligence (AI), or more precisely machine learning that uses so-called convolutional neural networks (CNNs; or simply nets). The article presents the Video Reuse Detector (VRD), a tool programmed by developer Tomas Skotare, that uses machine learning to identify visual similarities within a given audio-visual database (such as the SF archive). Third, we perform a practical test of our suggested methods by computationally comparing 'Letters from Stockholm' with four archival films from the SF archive: SF2061A (Anon. 1909a), SF2061B (Anon. 1909b), SF2063 (Anon. 1909c) and SF2066 (Anon. 1909d). These SF titles all contain film footage of Stockholm in 1909 shot by Swedish Biograph, sequences that Sahlberg copied, edited and recompiled in his 1964 documentary film. Our experimental algorithmic research has been developed at Humlab, Umeå University, within the EU-funded project European History Reloaded: Curation and Appropriation of Digital Audiovisual Heritage. This research project combines state of the art tracing and tracking technologies and ethnographic fieldwork to analyse how different forms of reuse and curation shape the appropriation of the digitized moving image heritage (Waysdorf 2021: n.pag.).

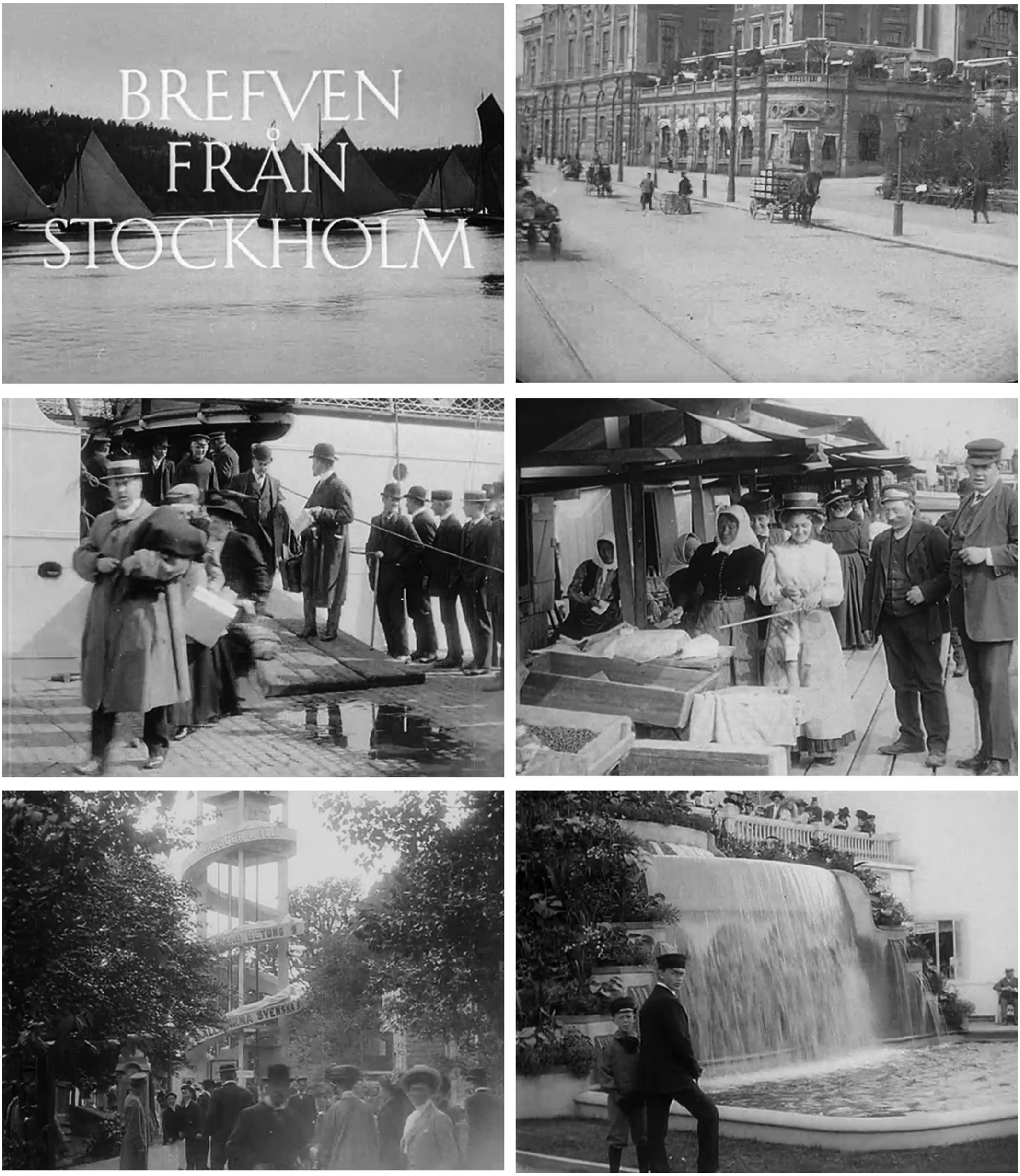
## ARCHIVAL GOLDMINES: THE SF ARCHIVE REVISITED

If televisual reuse was the major reason behind the acquisition of the SF archive in 1964, it needs to be stressed that SF had previously used its archive in a similar fashion. Initially, the company employed a film archivist, Knut Lindeberg, to keep track of film copies and registry cards. Later, during the 1950s, Sahlberg, employed by SF between 1941 and 1964 as script and songwriter, director and film historian, took an increased interest in the SF archive. In 1959 he rediscovered a number of early fiction films, such as Carl Engdahl's *Bröllopet på Ulfåsa* ('The wedding at Ulfåsa') from 1910. Foremost, however, he began working with the SF archive as a commercial asset for documentary film productions; archival reuse was his guiding concept (Sahlberg 1959: n.pag.). But the state of the archive posed problems. According to a later account by Sahlberg, Lindeberg had not been an archivist by profession and documentation had suffered: 'against his archival methodology a great deal could be said' (Sahlberg 1961: 38). Still, in 1961 Sahlberg edited and reused the oldest material in the SF archive, from 1897 to 1914, and produced the documentary film *När seklet var ungt* ('When the century was young'). This compilation film – *klippfilm* was the term used – received fine reviews and became rather popular at cinemas in Sweden. SF was praised for daring to experiment with and reuse old newsreels in novel ways, and Sahlberg was admired for having scrutinized at least '50,000 meters of film' ('Jerome' 1961: n.pag.). Then again, it was not a coincidence that this type of documentary film premiered at the time: 'The age of newsreels as contemporary depictions has ended – replaced by television' (Tjerneld 1961: n.pag.).

Inspired by his success, Sahlberg would reuse the SF archive repeatedly during the early 1960s, producing five compilation films for commercial release by SF prior to SR's purchase of the archive: *Hotad idyll: En film om hösten 1914* ('Threatened idyll: A film about the autumn of 1914') and *Kungen i närbild* ('The King in close-up'), both in 1962; *Fåfängans marknad* ('Vanity fair') in 1963; 'Letters from Stockholm' in 1964; *Klassiska svenska filmer* ('Classic Swedish films'), released in 1965 by SF (after the film archive was sold). These films were hardly commercial successes; still, they were important for the cultural prestige of SF. Sahlberg's films were screened as supporting features together with arthouse productions: 'Vanity fair' conjointly with Alf Kjellin's *Siska* (1962), 'Threatened idyll' together with Ingmar Bergman's *Tystnaden* (*The Silence*) (1963) and 'Letters from Stockholm' prior to Bergman's first colour film, the comedy *För att inte tala om alla dessa kvinnor* (*All These Women*) (1964). Most of Sahlberg's compilation films were reviewed in the daily press and received public attention, a fact that most likely was important to Radio Sweden's decision to purchase the SF archive.

In an article in *Svenska Dagbladet* in early 1963 about the forthcoming agreement between SF and SR – aptly entitled 'SF archive, a goldmine for TV' – Sahlberg was described as a man who 'with a loving hand managed the archive' (Anon. 1963: n.pag.). Consequently, after the deal was settled, it came as no surprise that SR hired him to work with the SF archive in a new media setting. In the late 1960s, Radio Sweden made an exceptional contribution to the cultural heritage in preserving the SF archive, migrating it to polyester-based safety film and cataloguing the material. At the same time, Sahlberg started producing TV programmes. In a memorandum from 1962 he had suggested that at least 130 TV programmes could be compiled from the SF archive; ideas were plentiful (Sahlberg 1962a). Sahlberg came to produce a range of TV programmes, all less well known than his earlier compilation films. They appeared regularly in the programme schedule; in November 1966, *Med ångbåt och ångtåg mot Östergyllen* ('With steamboat and steam train towards Östergyllen') was broadcast, described as a 'programme in the form of a Nils Holgersson trip through Sweden with the help of old newsreels from the rich SF archive that Radio Sweden now owns, managed by Gardar Sahlberg' (Anon. 1966: n.pag.). In a similar way, Sahlberg reused the SF archive in subsequent TV programmes such as *Bergslagsbygd och norrlandskust* ('The Bergslagen countryside and the coast of Norrland') in 1967 and *Stockholm under 20-talet* ('Stockholm during the twenties'), a TV programme in two parts broadcast in 1968. Sahlberg retired in the mid-1970s, but in years to come, content from the SF archive was repeatedly reused, becoming a recurrent feature in thousands of TV programmes (Snickars 2008).

There is an abundance of information about the prolific Gardar Sahlberg in the Swedish Film Database. It contains, for example, metadata about so-called 'Related films', listing some film titles that Sahlberg reused in




*Figures 1–6: The documentary film 'Letters from Stockholm: A film about the summer of 1909' was produced by SF and compiled by Gardar Sahlberg in 1964. Almost all sequences in the film were based on archival film footage from the SF archive shot in 1909, for example SF2061A, with street and harbour sequences from the city of Stockholm filmed by Swedish Biograph, or sequences from the Stockholm Exhibition of 1909, SF2063, also shot by Swedish Biograph. Used by permission.*

his compilation films. Sequences from the first Swedish film, *Konungens af Siam landstigning vid Logårdstrappan* ('The King of Siam's disembarkation at Logårdstrappan') from 1897 appeared in 'When the century was young', as did scenes from *Ingeborg Holm* (Sjöström 1913), and in 'Vanity fair' Sahlberg reused scenes from Urban Gad's early silent classic, *Afgrunden* (*The Abyss*) from 1910. Yet, the database information on previous titles is scant; only a few reused film titles appear, and almost nothing is said about newsreels or other types of footage from the SF archive. The metadata hardly explains how the reuse of older material was a fundamental practice of Sahlberg's filmmaking. If the daily press was correct that of 50,000 m of film, some 2000 m were reused in a single production, one can estimate that he reused 4–5 per cent of relevant content from the SF archive in a compilation film or TV programme. Today, sitting at a Steenbeck editing table (which Sahlberg used at SR), it would indeed be possible to manually detect this filmic reuse, even if such research would be demanding and time-consuming. In a digital setting – the SF archive has been openly available at SLBA and the National Library of Sweden since the early 2000s in the collection Journal Digital – it is easier, but still labour-intensive. Yet, by applying machine learning algorithms from the field of computer vision, the task can be reduced to (very rapid) calculations.

## FILM ARCHIVES AND COMPUTER VISION

It might seem a coincidence, but at the same time Gardar Sahlberg was compiling and reusing the SF archive at SR, a group of computer scientists at the Massachusetts Institute of Technology (MIT) gathered for the 1966 Summer Vision Project. Among them was Marvin Minsky, sometimes referred to as 'the father of artificial intelligence'. The task was to study how computers might understand images. A first step was to automatically divide pictures into regions based on surface and shape properties; these might then act as the basis for object identification (see Szeliski 2011). During the 1970s, computer vision began to be able to sort simple objects into categories (such as circles and squares); developments also included the ability to interpret typed text via optical character recognition (OCR). Gradually, computer vision became a subfield of AI, with the general purpose of training computers to process and interpret the visual world. Consequently, it has been argued that the idea to include images within the framework of AI 'marks a notable transition in media history, a moment when researchers began integrating image processing into the development of artificial intelligence, including the training of computers to read, detect, and describe aspects of pictures and visual environments' (Sayers 2019: 35).

To detect the reuse of archival material, computers – or rather computer software – need to be trained. On the one hand, such training is mundane. As online users we participate in computer vision tutorials almost every day when we encounter CAPTCHAs (Completely Automated Public Turing test to tell Computers and Humans Apart) and are asked to re-type distorted letters or identify cars, traffic lights and bicycles in street photographs. CAPTCHAs are a common way to distinguish between bots and humans, which at the same time repeatedly helps computers to interpret images as text. On the other hand, visual machine learning is a complicated task, and a reason why the digital humanities have focused on moving image material to a much lesser extent than text. But machines' ability to comprehend images, video and audio has progressed significantly in recent years (Fickers et al. 2018; Burghart et al. 2020) with the development of algorithms that can detect and analyse visual tropes in images, identify photographic techniques and track individual artistic styles of filmmakers (Champion 2017).

Then again, it is commercial actors and computer scientists, rather than humanists and media scholars, who have showcased the most interest in the analysis of digitized audio-visual data. While companies such as Google, Spotify and YouTube over time have processed millions of images, songs and videos for commercial purposes, similar computer vision algorithms are only now beginning to be used in the humanities (Arnold and Tilton 2020). In recent years, scholars have shown that visual computational methods can be productively applied to cultural heritage datasets to answer research questions concerning media history (Williams 2016; Jofre et al. 2020; Pustu-Iren et al. 2020). Cultural analytics, with Lev Manovich as its foremost proponent, is perhaps the best-known example of a novel research field dedicated to the computational study of culture, focusing particularly on visual media as massive datasets. Cultural analytics also acts as a link between the humanities and computer science:

> In computer science fields called computer vision and multimedia computing, researchers for many years have been publishing new algorithms for automatic detection of image content, artistic styles, photographic techniques, genres of TV and video, and applying them to progressively larger datasets.
>
> (Manovich 2020: 11)

It should be noted that cultural analytics in Manovich's terms is more prone to large-scale analyses of current visual media; the term refers primarily 'to computational analysis of patterns and trends in contemporary digital culture (as opposed to only historical culture)' (Manovich 2020: 7).

In general, the emergence of algorithmic ways of using and studying media archives is a recent phenomenon. For example, the 2017 Special Issue of this journal on film archival practices (*JSCA* 7:2) did not feature any articles or signs of this development (Brunow and Stigsdotter 2017: 75–78) – apart from an article co-authored by one of us (Norén and Snickars 2017: 155–75.). Yet at present a range of media archival undertakings and projects are devoted to the topic. Archivists, researchers and artists have shown increased interest in computer vision algorithms committed to archival reuse, with a lot of relevant code published on Github, an open repository for distributing source code. The video artist and YouTuber Denis Shiryaev, for example, makes use of publicly available AI-technology to upgrade black and white, grainy and low-resolution early cinema footage. One of his videos on YouTube, 'upscaled with neural networks' (Shiryaev 2021: n.pag.), reuses footage shot by Svenska Bio (Swedish Biograph) in New York in 1911. At the time this company was becoming the leading producer of film in Sweden, and the intention of shooting on location in New York was to film local scenes to be reused in a number of pre-planned fiction films (Snickars 2001). To increase footage resolution, Shiryaev has used Super-Resolution Generative Adversarial Network (SRGAN) algorithms that increase image sharpness, and Depth-Aware video frame INterpolation (DAIN) algorithms that can boost old footage up to 60 frames per second. Primarily, however, Shiryaev has algorithmically colourized the footage by using DeOldify, a black and

*Figures 7 and 8: Footage from New York in 1911 shot by Swedish Biograph, reused and algorithmically manipulated by video artist and YouTuber Denis Shiryaev. Used by permission.*

white image colourizing library. The result of this online video enhancement is transformative – or as *Wired* put it: 'AI magic makes century-old films look new' (Simon 2020: n.pag.).

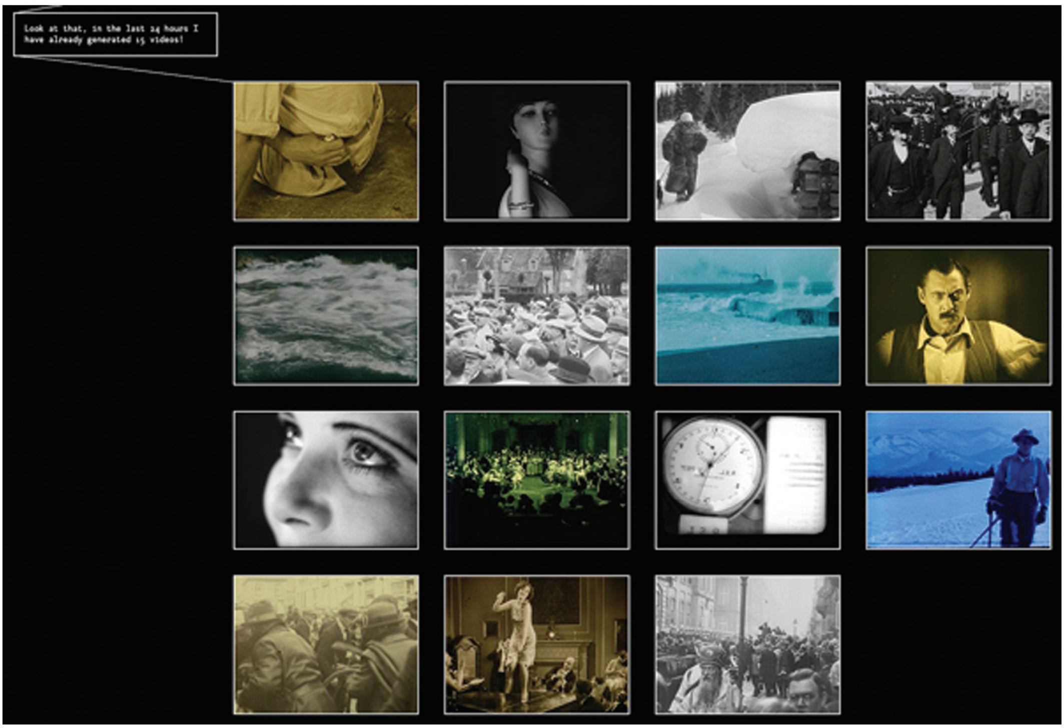
Shiryaev's algorithmic manipulation of the 1911 Swedish Biograph footage has been viewed almost 20 million times on YouTube. If Sahlberg's TV programmes based on the SF archive reached a large audience via television, platforms such as YouTube can encourage archival reuse and circulation today. Shiryaev's videos are thus an illustrative example of film heritage 'gone rogue', in the words of Abigail de Kosnik. Online, she argues, 'each media commodity becomes, at the instant of its release, an archive to be plundered, an original to be memorized, copied, and manipulated – a starting point or springboard for receivers' creativity, rather than an end unto itself' (de Kosnik 2016: 4). Shiryaev's videos are beautiful and compelling, film archivist Rick Prelinger has stated, 'but they show you something that never was. They're not archival; they're fiction' (Prelinger 2021: n.pag.). But Shiryaev has openly declared that his videos are indeed algorithmically manipulated; they have nothing in common with established film restoration practices. Still, in years to come AI manipulation may make tracing the origins of digital moving images more difficult.

If doubts have been raised as to archival 'authenticity [being] threatened by AI' (Prelinger 2021: n.pag.), other parts of the media archival sector adjacent to the content industry have a more affirmative view on the matter. Archival content can be monetized through AI. A webinar arranged by Quantiphi, an applied AI and data science software company, for example, was aptly entitled 'Finding gold in your media archives: Leveraging AI for content monetization' (Quantiphi 2021). It comes as no surprise that many major media archives have started using AI. The BBC employs machine-generated automatic speech recognition (ASR) to expand and refine its catalogue metadata; technologies that analyse audio, and especially speech, are significantly ahead of visual tools. Similarly, the Library of Congress recently launched Citizen DJ, a platform for users to create hip hop music by reusing the library's audio and video collections that have been annotated by

AI. Another sign of the times is the appropriately entitled article, 'Artificial intelligence: An object of desire', published by the International Federation of Television Archives (Bazán-Gil 2020). Moreover, the Netherlands Institute for Sound and Vision argued in a series of blog posts during the summer of 2021 that AI 'has enormous potential to expand the reach of audiovisual archives' (Cecchine 2021c: n.pag.). AI tools are being used for metadata creation (Cecchine 2021a), even if many audio-visual archives are still at the beginning 'of the process of formulating an approach to AI' (Cecchine 2021b: n.pag.).

In short, AI is seen as vital for the media archival sector, as is also apparent in the range of scholarly research projects concerned with new ways of algorithmically exploring media archives. In the United States, the Media Ecology Project at Dartmouth College has prototyped machine vision search for identifying objects and actions in archival video; at the University of Southern California, Virginia Kuhn runs the VAT – Video Analysis Tableau – Project applying computational analysis to the study of moving image media; at the University of Richmond, the Distant Viewing Lab uses computational techniques to analyse visual culture on a large scale. Similar research projects are also present in Europe; the Finnish project Methods for Managing Audiovisual Data (MeMAD) has provided novel methods for efficient reuse and re-purpose of multilingual audio-visual content (MeMAD 2021), while the project Visual Information Retrieval in Video Archives (VIVA) is aimed at supporting media archives by developing 'videomining software' to create 'new opportunities for information retrieval concerning scenes, people and similar images in so far undocumented material' (VIVA 2021: n.pag.). Of particular relevance for this article are research and curatorial activities at the EYE Film Museum in Amsterdam. To foster algorithmic reuse, in 2018 the museum acquired a filmmaking bot. With the mission to bring 'film heritage to the algorithmic age', Jan Bot, as the software programme is called, automatically produces short videos via reusage of 'a 100-year-old film archive, taking inspiration from today's trending topics' (Jan Bot 2021: n.pag.). Jan Bot uses computer vision technologies to annotate each shot in the found footage collection Bits & Pieces with a set of tags and then proceeds by using Natural Language Processing tools and Google Trends to select relevant footage for trending news. Jan Bot can be followed on Twitter (@janbotfilms).

The EYE Film Museum has also been the host for the research project Sensory Moving Image Archive: Boosting Creative Reuse for Artistic Practice and Research (SEMIA). The project, which ended in 2020, was devoted to ways in which the field of computer vision can contribute to novel forms of access to moving image collections. SEMIA set out to explore 'the affordances of deep learning techniques for revealing similarity-based patterns in (large) collections of digitised moving images', including an investigation of how visual analysis tools might help and foster 'more exploratory forms of engaging with digital archives' (Fossati et al. 2020: n.pag.). SEMIA was interested in the preconditions of archival reuse and tried to reorient visual analysis methods by drawing on artistic practices of archival moving image appropriation. An interest in new forms of algorithmic reuse lies at the core of several recent film archival projects. Archival remediation and the possibilities of novel ways of filmic appropriation in a new medium is thus what unites Shiryaev, Jan Bot and SEMIA with Gardar Sahlberg.

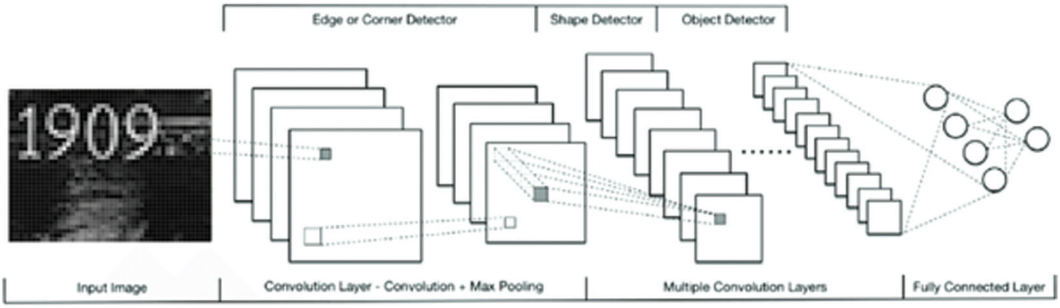

*Figures 9 and 10: The EYE Film Museum in the Netherlands features the film archival Jan Bot, a programmed piece of software dedicated to bringing film heritage into the algorithmic age. Based on contemporary trending themes – in this case the UEFA European Football Championship during the summer of 2021 – the bot selects relevant film material from the museum's collection of found footage and produces a short compilation film. Used by permission.*

## CNNS AND VIDEO REUSE

'How do computers see?', Melvin Wevers and Thomas Smits recently asked themselves in an article on the agency of computer vision models as 'optical instruments' (Wevers and Smits 2021: 331). In another text they have discussed how so-called CNNs can be used to categorize and analyse digitized historical visual sources (Wevers and Smits 2020). CNNs represent one of the latest developments in computer vision and provide a powerful method for analysing visual content with the help of algorithms. Put simply, CNNs constitute an AI or machine learning technique for analysing visual features in images. CNNs are trained to autonomously search for patterns of similarity in visual content by mapping, remembering and comparing the distribution and structure of pixels across digital images. After being taught to observe and interpret visual content, generally by training on large archives of images, the performance of CNNs is tested against benchmark datasets tagged and annotated by humans. This makes it possible to evaluate their ability to interpret footage. Currently, CNNs are used to examine everything from X-ray plates in medicine to faces in images posted on social media. As Wevers and Smits suggest, however, the technique can also be used to study historical images, primarily because of its capacity to quickly process large amounts of data (Wevers and Smits 2020).

Wevers and Smits have applied a CNN to millions of historical images collected from Dutch digitized books, newspapers, periodicals and historical documents. Their goal is primarily to make content – the actual visual features, not the metadata – of historical footage searchable. This is done by allowing a CNN to semantically annotate and interpret what appears in images, thereby making it possible to search for and retrieve footage using topics or keywords and/or specific image types (such as a photograph). With the help of this technique, Wevers and Smits map and identify visual trends in digitized newspapers, including transformations in the use of original illustrations and photography over time. They also improve techniques for navigating large visual databases by clustering similar-looking content (such as advertisements) and textually requesting the retrieval of images containing specific categories of objects (see Wevers and Smits 2020).

In our explorations of the SF archive, we apply a similar approach to study reuse. In particular, we set up a system for detecting if similar or identical video content re-emerges within selected videos from the SF archive, using a combination of a pre-trained CNN called ResNet50 and the so-called Facebook AI Similarity Search (FAISS) index, an open-source library for efficient similarity searches. Unlike Wevers and Smits, who start with still images, we begin our analysis with moving images, which adds a temporal dimension to our similarity searches. Furthermore, the starting point for our similarity search is a visual copy of film footage, as opposed to a textual string or keyword. As a result, our method is primarily geared towards tracing the re-emergence of a *specific* video, partial video clip or set of videos within an archive. Our focus is hence a highly designated footage depicting a specific city street using a specific camera and angle, not video clips depicting city streets in general. This technique has similarities to commercial techniques for tracing copyright abuse online, such as YouTube's digital fingerprinting system Content ID, where the goal is not to automate the search and retrieval of films containing, say, explosion scenes in general, but finding instances when a specific explosion scene from a specific copyright-protected film has been (potentially illegally) reused.





*Figures 11 and 12: In a digital image or one frame in a video sequence, a pixel is the smallest item of information. The same pixels form the basis for CNNs. They start with extracting a so-called feature map corresponding to the length and width of an image in pixels, then proceed algorithmically by looking at the pixels around a given pixel to identify an edge, corner or texture both horizontally and vertically to detect a distinctive shape, one pixel at a time. CNNs thus start with an image's raw pixel data (as input) and learn how to extract different features, one (very fast) step at a time, to ultimately infer and detect what object they constitute. Used by permission.*

We begin our analysis of the SF archive by dividing each film, or video in our terminology, into still frames. Commonly, a digital video consists of 24–30 frames per second, and we extract one such frame, or thumbnail, for every second of video. These thumbnails later constitute the main units of analysis when our SF videos are matched and compared. After saving the thumbnails, we apply an open-source CNN called ResNet50 to process the images and identify patterns of similarity across them. ResNet50 was developed by researchers at Microsoft and has been trained and validated using the benchmark dataset ImageNet (ImageNet Large Scale Visual Recognition Challenge), arguably the world's most frequently used method for developing and testing computer vision tools (ImageNet 2021).

Like all CNNs, ResNet50 performs its image processing in multiple layers of analysis and abstraction. Each layer in a CNN processes an input and produces an output, which is then passed on to the next layer. One layer in a CNN may, for instance, observe how pixels are spatially arranged and search for areas with a high contrast between nearby pixels – a good marker for what is visually unique in a particular image – while another layer might focus on reducing what information is stored about pixel contrasts, instructing the model to forget all areas in a picture with a lower pixel contrast than a given value. This way, the CNN produces a successively smaller and (hopefully) more precise map of the analysed image. Somewhere before the top layers of the analysis are reached, a CNN produces a compressed and final interpretation of the key visual characteristics of an image. It is then common for the final layers of a CNN to be trained to annotate or classify the content in images by estimating what objects appear in the footage.

The ResNet50 neural net is 50 layers deep and in our use of it, we instructed the model to stop its image processing at a satisfactory result (a compromise between the speed of data processing, the size of extracted data and the detail with which features were obtained). This allowed us to acquire a highly compressed visual interpretation of each thumbnail, while cutting out the common – and final – task of semantically interpreting, classifying and annotating what appeared in the images. Avoiding the final step simplifies and speeds up image processing. It is also sufficient for our purposes here, since we are solely focused on studying video reuse, as opposed to obtaining algorithmic, semantic interpretations of the content within our footage. In the next step, we index the compressed thumbnail interpretations using FAISS. This tool, developed by Facebook, helps store and index the compressed thumbnails as a set of vectors, that is, numerical representations of the analysed files (FAISS 2021). This further reduces their size and makes it possible to quickly compare and match different thumbnails against each other. FAISS was first released in 2019, and while the indexation method can be used for any sort of similarity search, we have used it to identify similarities between images. By combining ResNet50 and the FAISS library, we then calculate a so-called distance metric for each matched pair of thumbnails, a value that is an estimation of the visually most similar 'neighbours' for every extracted frame. The distance metric is indicative of similarity, with low values (short distances) indicating high visual similarity and high values (long distances) indicating low visual similarity. Finally, we create a system for filtering and narrowing down the search results, as well as obtaining the top matching results in the form of both textual lists and visual representations. We call this method for combining CNNs, indexation libraries, filtering techniques and visualization tools the VRD toolkit.

To summarize, the applied method can be described in the following seven steps: (1) extract one image thumbnail per second of the analysed video; (2) process thumbnails using the neural network ResNet50 to extract their key visual characteristics; (3) stop the image processing when ResNet50 has reached a satisfactory result and store the compressed thumbnail interpretations as separate files; (4) index the compressed thumbnail interpretations using the FAISS library; (5) calculate distance metric for each matched pair of frames, thus locating the visually most similar neighbours for every analysed thumbnail; (6) filter the matching results according to desired settings; (7) output list of the top matching results, as well as visual examples and illustrations of image similarities.

## VIEWING 'LETTERS FROM STOCKHOLM' WITH NEURAL NETS

'Letters from Stockholm' premiered in the summer of 1964. The film recounts a fictional visit to Stockholm in 1909. Gardar Sahlberg's script takes the form of letters from a woman to an unknown recipient; Britta Holmberg acted as narrator. The film begins on a steamboat from the Stockholm archipelago arriving at Skeppsbron, the long quay in Gamla stan (Old Town). In about eighteen minutes, the film takes the viewer on a cinematic trip through Stockholm, travelling on a tram and seeing various sights of the capital. The reused footage consists mainly of street scenes of Stockholm shot by Swedish Biograph; harbour sequences and scenes from the Stockholm Exhibition of 1909 were followed by a ride upwards in the public lift Katarinahissen (constructed in 1881), where Sahlberg reused a short sequence from the Pathé film *Visite à Stockholm* with a characteristic binocular mask. This film was from 1907, but Sahlberg included it anyway. On a few occasions, Sahlberg inserted other types of material, close-ups of handwritten letters from the narrator or a menu from Hotel Rydberg. Diegetic sound as well as music was added to the silent footage. But overall, the preserved films in the SF archive prescribed what Sahlberg could show; the archive itself was the prerequisite for his compilation film. 'Letters from Stockholm' received praise in the daily press: 'a delightful little film', *Svenska Dagbladet* stated, 'for many it was probably the real treat of the evening' rather than Bergman's comedy in colour (*All These Women*) (Unger 1964: n.pag.). The film critic Mauritz Edström was very pleased: 'This time, Sahlberg has found exactly the right grip on his material […] It is a pure pleasure to follow [him] on this cultural historical journey through a living past. The short film alone is worth the ticket price' (Edström 1964: n.pag.).

To detect exactly what archival film footage from the SF archive Sahlberg reused in his compilation film, we decided to compare it with four films from the archive: SF2061A, SF2061B, SF2063 and SF2066. These films all include footage from Stockholm in 1909; they are all about ten minutes long, and probably it was SF's first film archivist, Knut Lindeberg, who gave them numerical titles. Sahlberg included other types of footage from the SF archive in his film, but for the sake of pedagogical clarity we will focus on these four SF titles. To illustrate how Sahlberg's practices of reuse can be unpacked with archival algorithms, we ran the VRD to trace how and when footage from the four archival films reappeared in 'Letters from Stockholm'. In computational terminology, we thus perceive Sahlberg's compilation film as one dataset that we have compared with four other datasets, the four SF titles. To make things somewhat complicated, the copy of Sahlberg's film that we used in our tests, taken from the database Journal Digital at the National Library of Sweden, is divided into two parts; hence the file reference to 'Letters from Stockholm Part 1'. One also needs to keep in mind that the VRD tool compares video frames rather than sequences; for each second in all our datasets, one frame was extracted, resulting in thousands of video frames.

The (still) rudimentary interface of the VRD lets the film historian compare different archival footage based on image similarity. As stated, the VRD tool supplies a distance metric for each match that estimates the visually most similar neighbours for every analysed video frame within a dataset. A low distance metric indicates high similarity, and conversely, the higher the value, the lesser the likeness. Since (in theory) an infinite number of similarity matches between video frames can be detected, we concentrate on four cases as illustrative examples of matches. The purpose is to demonstrate how the VRD tool functions and operates; we will discuss both its advantages and shortcomings.

Our first example is from a video frame taken some six minutes into Sahlberg's compilation film: to be exact, the archival reference frame 357.png (in red) was extracted 357 seconds into 'Letters from Stockholm'. For this particular frame, the VRD tool suggests three similarity matches. The first of these, frame 4.png from SF2061A (in blue), is more or less identical to the reference frame 357, with a distance metric of 18,766.45. The frame displays the Vasa Bridge in central Stockholm, with the (in 1909) brand new tram track. Both frames, 357.png and 4.png, are extracted from a short sequence filmed by Swedish Biograph from a moving tram travelling across the bridge. The sequence is originally only four seconds long, and in this case one frame extraction is also enough for scene comparison. Second, according to the VRD tool, there exists a certain similarity to frame 457.png (in green) taken from SF2066, with a distance metric of 56,461.30. To a human, however, it is obvious that this video frame does not display the actual Vasa Bridge. Then again, the comparison reveals some visual likeness, an empty space in the foreground and a building in the background of the frame. Finally, the VRD tool suggests yet another match (in yellow), with frame 445.png from SF2061B. To a human eye this frame is a more similar neighbour to the archival reference frame – but not to the VRD tool, since the distance metric is higher than for the previous frame. This video frame is also taken from a sequence shot from a tram travelling over a bridge, but this time over another bridge, Norrbro, towards Gustav Adolfs torg (square), with the Swedish parliament building to the right (see Figure 13).

Our next example moves from bridges to buildings and displays similarities that the VRD tool detects in sequences shot in front of the Royal Swedish Opera in Stockholm. Towards the end of the first part of Sahlberg's film there

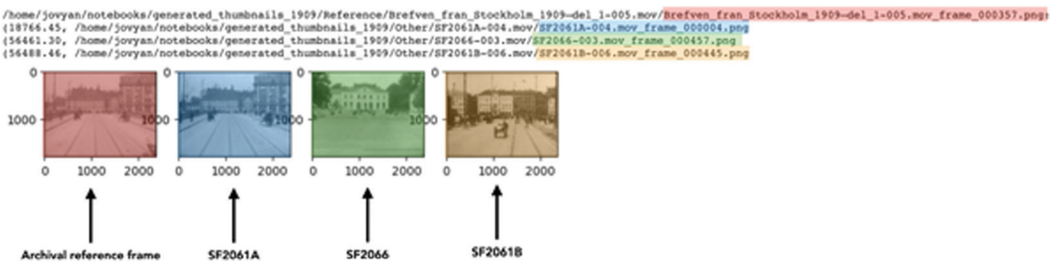



*Figure 13: The VRD, developed at Humlab, Umeå University, lets the film historian compare archival footage based on image similarity. In this case, the archival reference frame (in red) from Sahlberg's 1964 compilation film is almost identical to the frame from SF2061A (in blue) of the Vasa Bridge in central Stockholm. The frame extracted from SF2061B (in yellow), however, displays the wrong bridge (Norrbro). Used by permission.*

is a panoramic sequence, shot from a tram slowing down in front of the Opera Terrace, one of the capital's most prestigious restaurants. The VRD tool suggests four similarity matches compared with the archival reference frame 567.png (in red). The first of these, frame 104.png (in blue), with a distance metric of 16,653.09, appears two minutes into SF2061A. Second, the VRD tool suggests that frame 320.png (in green) from SF2061B shares a similarity with the reference frame; the distance metric is set to 36,995.37. This comparison is particularly interesting since the static sequence in SF2061B clearly displays the Opera Terrace, even if the windy sequence of six seconds is filmed from a distance (compared to Sahlberg's film). The match is also indicative of the accuracy of the VRD tool; the Royal Swedish Opera indeed appears in frame 320.png, but not cinematically framed in exactly the same way as in frames 567.png and 104.png. As in our previous example, the VRD tool also displays two matches that for a human are obviously not identical to the archival reference frame. Both frame 454.png (in yellow) and frame 299.png (in pink) do feature a building and a fairly empty foreground, but they are clearly not scenes that Sahlberg reused in his film (see Figure 14).

Our third example examines crowds of people. Since 'Letters from Stockholm' featured a number of street scenes from the capital, there are plenty of Stockholmers visible in Sahlberg's film. Frame 169.png (in red) appears after some three minutes in his compilation film: 'At Mälartorget and Kornhamnstorg [two squares in the city], commerce was in full swing', as the narrator states. The VRD tool detects a similar sequence in SF2061A, frame 500.png (in blue), but it is the other three matches that are of more interest. They also display crowds of people, yet at completely different settings. In frame 114.png (in green) people are entering the 1909 Stockholm Exhibition; in frame 301.png (in yellow), taken from SF2066, travellers are seen disembarking from a steamboat and in frame 170.png (in pink), an audience watches a military parade marching pass the film camera. In all three frames (and sequences), crowds of people are evidently present, which the VRD tool detects. It accurately matches the commerce at Mälartorget and Kornhamnstorg in one frame, but otherwise fails to do so, rather suggesting an iconic likeness of crowds in the footage at hand, if one follows Charles Sanders Peirce's classes of signs denoting an object (icon, index and symbol) (Peirce 1935). To a human there is an evident resemblance between these five video frames, but the variety of the distance metric value clearly displays that the VRD tool does not consider

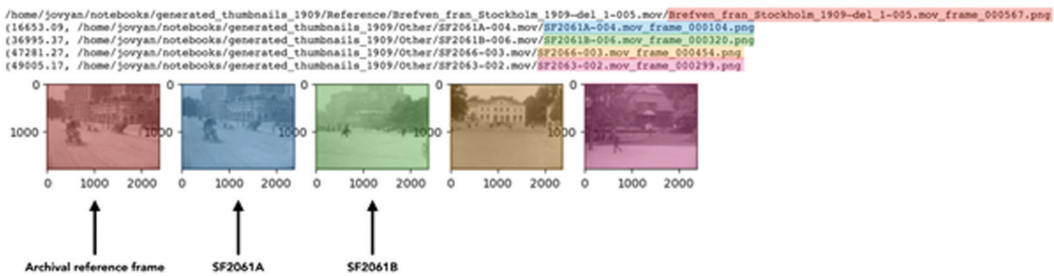



*Figure 14: The archival reference frame 567.png from Sahlberg's compilation film depicts the Royal Swedish Opera in central Stockholm; the VRD suggests four similar video frames. Two of these, frame 104 (in blue) and frame 320 (in green), do indeed display the same building, but the other two, frame 454 (in yellow) and frame 299 (in pink) do not. Used by permission.*

them particularly similar. For frame 500.png the value is 17,637.67. As with our other examples, a value below 20,000 usually suggests a more or less perfect match. The distance metric of the other three frames, however, is relatively high, more than 70,000, and as stated previously, the higher the value, the less the correspondence (see Figure 15).

In our final example we move from crowds to heights. Tom Gunning once described the 'view aesthetic' as an '*Urform* of early nonfiction film'; it highlighted the way 'early actuality films were structured around presenting something visually' (Gunning 1997: 14). In Sahlberg's compilation film there are a number of sequences with a similar view aesthetic, foremost filmed from the lift Katarinahissen. The archival reference frame 244.png (in red) is extracted about four minutes into Sahlberg's film, and the VRD tool suggests four similarity matches. The first, frame 441.png (in blue) from SF2061A, is indeed a similar neighbour, albeit shot at a slightly different angle. However, the next three matches, frame 25.png (in green), frame 296.png (in yellow) and frame 150.png (in pink), do not portray Katarinahissen at all. The distance metric of these video frames is relatively high, over 60,000. But on closer inspection all frames are somewhat similar to the Katarinahissen frame. Two of the frames, for example, are taken from establishing shots of Norrbro filmed from a high angle (in yellow) and of the 1909 exhibition (in pink). In other words, the VRD tool detects distance and a certain 'view aesthetic', even if more exact matches are lacking (see Figure 16).

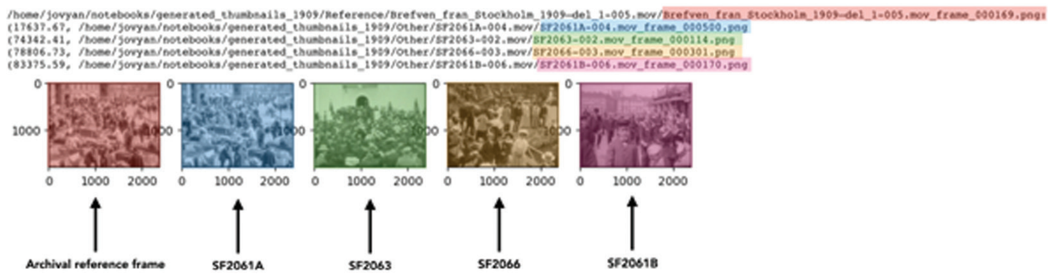



*Figure 15: The VRD tool detects sequences and video frames with people but has some problems with specific crowds. It accurately makes a similarity match between the archival reference and frame 500.png in SF2061A (blue) but is otherwise unsuccessful in computing exact similarities. Used by permission.*

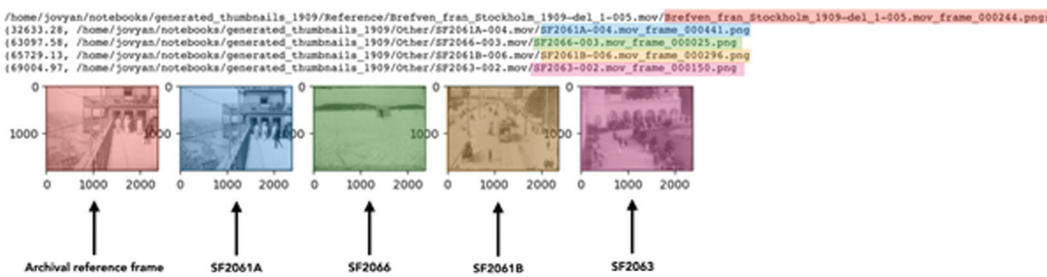



*Figure 16: If crowds pose a problem for the VRD tool, so do heights. Frame 441.png (in blue) is a similar neighbour to the archival reference frame 244.png (in red); they both depict the lift Katarinahissen in Stockholm. The other video frames display dissimilar places, but are also establishing shots, with two filmed from a high angle. Used by permission.*

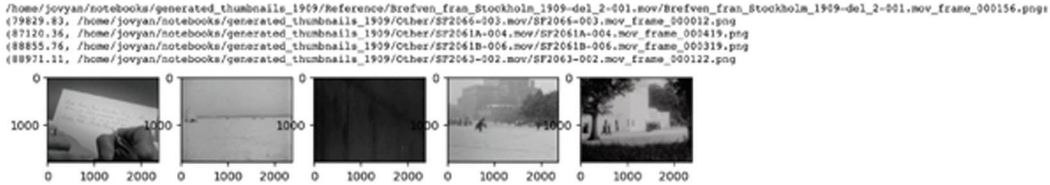



*Figure 17: VRD failure – no matches?*

## CONCLUSION

In the introduction to *Uncertain Archives: Critical Keywords for Big Data*, the editors argue that machine learning and AI are 'highly metaphorical notions'. They are furthermore associated with automatic 'gathering and management of big data', implying an 'independence of technology from the human'. While the 'digging in the uncertain archives of big data' indeed needs 'human-oriented perspectives' and critical reflections (Bonde Thylstrup et al. 2021: 1–2), scholars in the humanities also need to get their hands dirty by concretely working with machine learning and AI. They would then discover that machine learning is far from being only metaphorical. Training software is often a tiresome empirical endeavour, usually mundane and time-consuming, suggesting a constant interplay between developer and code, scholar and content, to produce accurate and valid output. Such work is imperative; the critical scholarly gaze needs to engage with AI and come to grips with what machine learning *does*. Our work with developing the VRD tool – based on machine learning that uses convolutional neural nets – is a very practical research undertaking.

Because Gardar Sahlberg's compilation film 'Letters from Stockholm' is eighteen minutes long, the VRD tool extracted more than 1000 video frames from it. In this article, we have repeatedly stressed that Sahlberg used the SF archive as his empirical source. Comparing reuse in his film with a number of SF titles would inevitably lead to matches of similar neighbours, since the SF archive de facto was the prerequisite for his film. The results that the VRD tool presented confirm this assumption. As is evident from our examples above, in almost all cases the archival reference frame from Sahlberg's film had a close similarity match with the selected archival footage, usually with a distant metric value below 20,000. This result is hardly surprising; frames, scenes and sequences are similar since Sahlberg actively copied and reused them from the SF archive. Consequently, the only instances when the VRD tool does not display similarities is when Sahlberg had inserted other materials in his film, such as short scenes with handwritten letters. These sequences, filmed in the early 1960s, were not part of the SF archive and therefore did not result in any close similarity matches.

In this article we have explored and tested how algorithms can help identify instances of audio-visual reuse in a large dataset. One result is that it is indeed possible to understand Sahlberg's compilation strategies with neural nets; our VRD tool has repeatedly proven itself capable of detecting similar video frames. Our test is small-scale, and naturally, a video frame extracted every second is not the same thing as a sequence; after all, compilation film is mostly about editing temporal scenes and sequences. The robustness and capacity of the VRD tool, however, suggests that it can be of use to media scholars or archivists interested in researching how audio-visual content is reused on a larger scale. In our article, the use case is very small, but the more far-reaching idea is that these methods, supported by the VRD tool, could be used in a larger context. It would be possible, for example, to analyse a larger part of the SF archive – say, all footage from 1897 to 1914 that Sahlberg reused in his documentary 'When the century was young' – thus gaining new knowledge of film history by computational means.

## FUNDING

This article is based on scholarly work within the research project European History Reloaded: Curation and Appropriation of Digital Audiovisual Heritage (CADEAH), funded by the European Union's Joint Programming Initiative in Cultural Heritage and Global Change, Digital Heritage.

## REFERENCES

Anon. (1963), 'SF:s journalarkiv guldgruva för TV', *Svenska Dagbladet*, 21 March, n.pag.

Anon. (1964), 'SF-journaler till TV för tre miljoner', *Dagens Nyheter*, 4 January, n.pag.

Anon. (1966), 'Punkt för punkt: tv-tablå', *Svenska Dagbladet*, 16 November, n.pag.

Arnold, Taylor and Tilton, Lauren (2020), 'Distant viewing toolkit: A python package for the analysis of visual culture', *Journal of Open Source Software*, 5:45, pp. 1–6.

Bazán-Gil, Virginia (2020), 'Artificial intelligence: An object of desire', International Federation of Television Archives, 13 May, https://fiatifta.org/artificial-intelligence-an-object-of-desire/. Accessed 6 October 2022.

Bonde Thylstrup, Nanna, Agostinho, Daniela, Ring, Annie, D'Ignazio, Catherine and Veel, Kristin (eds) (2021), *Uncertain Archives: Critical Keywords for Big Data*, Cambridge, MA: MIT Press.

Brunow, Dagmar (2017), 'Curating access to audiovisual heritage: Cultural memory and diversity in European film archives', *Image & Narrative*, 18:1, pp. 97–110.

Brunow, Dagmar and Stigsdotter, Ingrid (2017), 'Scandinavian cinema culture and archival practices: Collecting, curating and accessing moving image histories', *Journal of Scandinavian Cinema*, 7:2, pp. 75–78.

Burghart, Manuel, Heftberger, Adelheid, Pause, Johannes, Walkowski, Niels-Oliver and Zeppelzauer, Matthias (2020), 'Digital humanities & film studies: Analyzing the modalities of moving images', *Digital Humanities Quarterly*, 14:4, http://www.digitalhumanities.org/dhq/vol/14/4/index.html. Accessed 6 October 2022.

Cecchine, Randi (2021a), *AI in Audiovisual Archives: What Kinds of Analyses Are Possible*, Hilversum NL: Netherlands Institute for Sound and Vision, https://beeldengeluid.nl/en/knowledge/blog/ai-audiovisual-archives-what-kinds-analysis-are-possible. Accessed 6 October 2022.

Cecchine, Randi (2021b), *Bringing AI into the Audiovisual Archive*, Hilversum, NL: Netherlands Institute for Sound and Vision, https://beeldengeluid.nl/en/knowledge/blog/bringing-ai-audiovisual-archive. Accessed 6 October 2022.

Cecchine, Randi (2021c), *The Potential for AI in Audiovisual Archives*, Hilversum, NL: Netherlands Institute for Sound and Vision, https://beeldengeluid.nl/en/knowledge/blog/potential-ai-audiovisual-archives. Accessed 6 October 2022.

Champion, Erik Malcolm (2017), 'Digital humanities is text heavy, visualization light, and simulation poor', *Digital Scholarship in the Humanities*, 32:1, pp. 25–32.

Edström, Mauritz (1964), 'Vår hufvudstad: så olik sig och ändå lik', *Dagens Nyheter*, 21 June, n.pag.

Facebook AI Similarity Search (FAISS) (2021), https://ai.facebook.com/tools/faiss/. Accessed 6 October 2022.

Fant, Kenne (1963), letter to Harry Schein, 9 May, Stockholm: Swedish Labour Movement's Archives and Library, Harry Schein Archive, vol. 4/33, 'Handlingar rörande filmpolitik, Filmutredningar 1963–1974'.

Fickers, Andreas, Snickars, Pelle and Williams, Mark (eds) (2018), 'Editorial', *VIEW Journal of European Television History and Culture*, Special Issue: 'Audiovisual Data in Digital Humanities', 7:14, pp. 1–4.

Fossati, Giovanna, Masson, Eef, Gosvig Olesen, Christian and van Noord, Nanne (2020), 'Exploring digitised moving image collections: The SEMIA project, visual analysis and the turn to abstraction', *Digital Humanities Quarterly*, 14:4, http://www.digitalhumanities.org/dhq/vol/14/4/000497/000497.html. Accessed 6 October 2022.

Gunning, Tom (1997), 'Before documentary: Early nonfiction films and the "view" aesthetic', in D. Hertogs and N. de Klerk (eds), *Uncharted Territory: Essays on Early Nonfiction Film*, Amsterdam: Nederlands Filmmuseum, pp. 9–24.

ImageNet (2021), https://image-net.org. Accessed 6 October 2022.

Jan Bot (2021), 'Jan Bot', https://www.jan.bot/about. Accessed 6 October 2022.

'Jerome' (1961), 'När seklet var ungt', *Dagens Nyheter*, 10 January, n.pag.

Jofre, Ana, Cole, Josh, Berardi, Vincent, Bennett, Carl and Reale, Michael (2020), 'What's in a face? Gender representation of faces in time, 1940s–1990s', *Journal of Cultural Analytics*, 5:1, 19 March, https://culturalanalytics.org/article/12266-what-s-in-a-face-gender-representation-of-faces-in-time-1940s-1990s. Accessed 6 October 2022.

Kosnik, Abigail de (2016), *Rogue Archives: Digital Cultural Memory and Media Fandom*, Cambridge, MA: MIT Press.

Manovich, Lev (2020), *Cultural Analytics*, Cambridge, MA: MIT Press.

Methods for Managing Audiovisual Data (MeMAD) (2021), https://memad.eu/about/. Accessed 6 October 2022.

Norén, Fredrik and Snickars, Pelle (2017), 'Distant reading the history of Swedish film politics in 4500 governmental SOU reports', *Journal of Scandinavian Cinema*, 7:2, pp. 155–75.

Peirce, Charles Sanders (1935), *The Collected Papers of Charles S. Peirce*, vol. 2, Cambridge, MA: Harvard University Press.

Prelinger, Rick (2021), 'NFTs and AI are unsettling the very concept of history', *Wired*, 20 April, n.pag.

Pustu-Iren, Kader, Sittel, Julian, Mauer, Roman, Bulgakowa, Oksana and Ewerth, Ralph (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. Accessed 6 October 2022.

Quantiphi (2021), 'Finding gold in your media archives: Leveraging AI for content monetization', webinar, https://pages.awscloud.com/apn-tv-641.html?trk=apn-tv-carousel. Accessed 6 October 2022.

Sahlberg, Gardar (1959), 'Från filmens barndom', *Svenska Dagbladet*, 19 December, n.pag.

Sahlberg, Gardar (1961), 'När seklet var ungt', *Chaplin*, 17 February, pp. 38–39.
Sahlberg, Gardar (1962a), 'PM angående Svensk Filmindustris journalarkiv 1962', Radio Sweden, Stockholm: Förvaltnings- och dokumentarkiv 'F2:3 Sahlberg'.
Sayers, Jenterey (2019), 'Bringing Trouvé to light: Speculative computer vision and media history', in K. Kee and T. Compeau (eds), *Seeing the Past with Computers*, Ann Arbor, MI: University of Michigan Press, pp. 32–59.
Simon, Matt (2020), 'AI magic makes century-old films look new', *Wired*, 8 December, n.pag.
Snickars, Pelle (2001), *Svensk film och visuell masskultur 1900*, Stockholm: Aura förlag.
Snickars, Pelle (2003), 'Digitalt filmarkiv under uppbyggnad', *Svenska Dagbladet*, 25 March, n.pag.
Snickars, Pelle (2006), 'Om utställningen som mediearkiv', in A. Ekström, S. Jülich and P. Snickars (eds), *1897: Mediehistorier från Stockholmsutställningen*, Stockholm: Mediehistoriskt arkiv, pp. 127–66.
Snickars, Pelle (2008), 'Televisionen som arkiv: om webb–tv', in A. Edin and P. Vesterlund (eds), *Svensk Television: en mediehistoria*, Stockholm: Mediehistoriskt arkiv, pp. 309–35.
Svensk filmdatabas (n.d.), https://www.svenskfilmdatabas.se. Accessed 7 October 2022.
Szeliski, Richard (2011), *Computer Vision: Algorithms and Applications*, London: Springer.
Tjerneld, Staffan (1961), 'Sensationell svensk filmkrönika från seklets början', *Expressen*, 10 January, n.pag.
Unger, Gunnar (1964), 'Sommarnattens grimaser', *Svenska Dagbladet*, 16 June, n.pag.
Vesterlund, Per (2012), 'Arkivvärde, programvärde, marknadsvärde', in M. Jönsson and P. Snickars (eds), *'Skosmörja eller arkivdokument?' Om filmarkivet.se och den digitala filmhistorien*, Stockholm: Mediehistoriskt arkiv, pp. 49–66.
Vesterlund, Per (2019), *Ur Harry Scheins arkiv*, Lund: Mediehistoriskt arkiv.
Visual Information Retrieval in Video Archives (VIVA) (2021), https://projects.tib.eu/en/viva/projekt/. Accessed 6 October 2022.
Waysdorf, Abby (2021), 'Remix in the age of ubiquitous remix', *Convergence*, 27:4, 3 March, pp. 1129–44, https://doi.org/10.1177/1354856521994454. Accessed 6 October 2022.
Wevers, Melvin and Smits, Thomas (2020), 'The visual digital turn: Using neural networks to study historical images', *Digital Scholarship in the Humanities*, 35:1, pp. 194–207.
Wevers, Melvin and Smits, Thomas (2021), 'The agency of computer vision models as optical instruments', *Visual Communication*, 21:2, pp. 329–49.
Williams, Mark (2016), 'Networking moving image history: Archives, scholars and the media ecology project', in C. Acland and E. Hoyt (eds), *The Arclight Guidebook to Media History and the Digital Humanities*, Falmer: Reframe Books, pp. 335–46.

## FILM AND TELEVISION REFERENCES

Anon. (1897), *Konungens af Siam landstigning vid Logårdstrappan*, Sweden.
Anon. (1907), *Visite à Stockholm*, France.
Anon. (1909a), *SF2061A*, Sweden.

Anon. (1909b), *SF2061B*, Sweden.
Anon. (1909c), *SF2063*, Sweden.
Anon. (1909d), *SF2066*, Sweden.
Bergman, Ingmar (1963), *Tystnaden*, Sweden.
Bergman, Ingmar (1964), *För att inte tala om alla dessa kvinnor*, Sweden.
Engdahl, Carl (1910), *Bröllopet på Ulfåsa*, Sweden.
Gad, Urban (1910), *Afgrunden*, Denmark.
Kjellin, Alf (1962), *Siska*, Sweden.
Sahlberg, Gardar (1962b), *Hotad idyll: En film om hösten 1914*, Sweden.
Sahlberg, Gardar (1962c), *Kungen i närbild*, Sweden.
Sahlberg, Gardar (1963), *Fåfängans marknad*, Sweden.
Sahlberg, Gardar (1964), *Brefven från Stockholm: En film om sommaren 1909*, Sweden.
Sahlberg, Gardar (1965), *Klassiska svenska filmer*, Sweden.
Sahlberg, Gardar (1966), *Med ångbåt och ångtåg mot Östergyllen*, Sweden: Sveriges Radio.
Sahlberg, Gardar (1967), *Bergslagsbygd och norrlandskust*, Sweden: Sveriges Radio.
Sahlberg, Gardar (1968), *Stockholm under 20-talet*, Sweden: Sveriges Radio.
Shiryaev, Denis (2021), *A Trip through New York City in 1911*, USA, https://www.youtube.com/watch?v=hZ1OgQL9_Cw. Accessed 6 October 2022.
Sjöström, Victor (1913), *Ingeborg Holm*, Sweden.

## SUGGESTED CITATION

Eriksson, Maria, Skotare, Tomas and Snickars, Pelle (2022), 'Understanding Gardar Sahlberg with neural nets: On algorithmic reuse of the Swedish SF archive', *Journal of Scandinavian Cinema*, 12:3, pp. 225–47, https://doi.org/10.1386/jsca_00075_1

## CONTRIBUTOR DETAILS

Maria Eriksson is a postdoctoral fellow at Humlab, Umeå University, and a visiting scholar at Basel University, Department of Art, Media and Philosophy. She studies the history and politics of everyday digital technologies. Her research is located at the intersection of software studies, media archaeology, social anthropology and science and technology studies and focuses on the interplay between culture and technology.

Contact: maria.c.eriksson@umu.se

https://orcid.org/0000-0002-7534-4268

Tomas Skotare is system developer at Humlab, Umeå University. He earned a Ph.D. in analytical chemistry and works as a developer and programmer in several digital humanities projects at Humlab, for example Welfare State Analytics: Text Mining and Modeling Swedish Politics, Media and Culture, 1945–1989.

Contact: tomas.skotare@umu.se

https://orcid.org/0000-0001-8445-0559

Pelle Snickars is a professor of digital culture at Lund University. He is a cinema scholar turned media historian working with digital methods. He is currently in charge of the research projects Welfare State Analytics: Text Mining and Modeling Swedish Politics, Media and Culture, 1945–1989 and Modern Times 1936, which explore how computational methods can help us understand modernity in new ways.

Contact: pelle.snickars@kultur.lu.se

https://orcid.org/0000-0001-5122-1549