

Fachartikel aus:

Hartmut Beyer / Thomas Mandl (Hg.): Bildähnlichkeit und Bildsuche: Geistes- und informationswissenschaftliche Zugänge zu historischem Material (= Zeitschrift für digitale Geisteswissenschaften / Sonderbände, 8). 2026. DOI: [10.17175/sb008](https://doi.org/10.17175/sb008)

Titel:

Bodies in Relation. Theater Photography and the Distant Viewing of Performance

Autor\*in:

Ulf Otto

Kontakt: [ulf.otto@lmu.de](mailto:ulf.otto@lmu.de)

Institution: Ludwig-Maximilians-Universität München

GND: [1032235454](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63864-p0011-9) ORCID: [0000-0002-3528-6972](https://orcid.org/0000-0002-3528-6972)

Contribution (CRediT): [Conceptualization](#) | [Writing – original draft](#) | [Methodology](#) | [Software](#)

---

DOI des Beitrags:

[10.17175/sb008\\_006](https://doi.org/10.17175/sb008_006)

Nachweis im OPAC der Herzog August Bibliothek:

[1934020788](#)

Erstveröffentlichung:

30.06.2026

Lizenz:

Sofern nicht anders angegeben 

Letzte Überprüfung aller Verweise:

12.09.2025

Format:

PDF ohne Paginierung, Lesefassung

GND-Verschlagwortung:

[Bildererkennung](#) | [Fotografie](#) | [Maschinelles Sehen](#) | [OpenCV](#) | [Theaterwissenschaft](#)

Empfohlene Zitierweise:

Ulf Otto: Bodies in Relation. Theater Photography and the Distant Viewing of Performance. In: Hartmut Beyer / Thomas Mandl (Hg.): Bildähnlichkeit und Bildsuche: Geistes- und informationswissenschaftliche Zugänge zu historischem Material (= Zeitschrift für digitale Geisteswissenschaften / Sonderbände, 8). Wolfenbüttel 2026. 30.06.2026. HTML / XML / PDF. DOI: [10.17175/sb008\\_006](https://doi.org/10.17175/sb008_006)

Ulf Otto

# Bodies in Relation. Theater Photography and the Distant Viewing of Performance

---

## Abstract

Theater photography created a massive visual archive of 20<sup>th</sup> century performing arts, largely neglected due to the medium's precarious status and archive size. This paper argues *deep learning* for *computer vision* can tackle both challenges, enabling *distant viewing* approaches from Digital Literary Studies and Art History applied to Theater Studies. Focusing on scene shots, it discusses why scale makes these materials interesting and what questions emerge. It addresses disciplinary objections, suggesting the approach's benefit lies in challenging prevailing paradigms. The paper reviews image composition, theatrical gaze, and technologies like *pose estimation*, *mesh recognition*, and *sentiment analysis*. It explores using biased industrial technologies as scientific instruments, arguing that scale provokes empirical outlook, shifting from aesthetic performance discussion to collective meanings through social mediation centered on the body.

Theaterfotografie schuf ein umfangreiches visuelles Archiv der darstellenden Künste des 20. Jahrhunderts, das wegen des prekären Status des Mediums und der Archivgröße vernachlässigt wurde. Der Aufsatz argumentiert, dass *Deep Learning* für *Computer Vision* beide Herausforderungen bewältigen kann und sich *Distant-Viewing*-Ansätze aus Digitaler Literaturwissenschaft und Kunstgeschichte auf Theaterwissenschaft anwenden lassen. Mit Fokus auf Szenenaufnahmen wird diskutiert, warum Skalierung diese Materialien interessant macht. Disziplinäre Einwände werden adressiert und es wird vorgeschlagen, dass der Nutzen des Ansatzes im Herausfordern vorherrschender Paradigmen liegt. Die Arbeit untersucht Bildkomposition, die theaterspezifische Perspektive und Technologien wie *Pose Estimation*, *Mesh Recognition* und *Sentiment Analysis*. Sie erforscht die Nutzung »voreingenommener« kommerzieller Technologien als wissenschaftliche Instrumente und argumentiert, dass Skalierung empirische Perspektiven provoziert. An die Stelle der ästhetischen Diskussion einzelner Aufführungsereignisse rückt die Untersuchung der medialen Repräsentation theatraler Körperbilder.

## 1. Introduction

A theater photograph is a strange entity. In a daily newspaper it is routine, an impression of what is currently going on. But once it becomes historical, once it loses its actuality, its grand gestures and lavish dresses become more alien with every day. The absence of voice, of action and often of color come to the fore. Detached from journalistic context, the image quickly loses meaning, and even when that is reconstructed, it stays questionable as to its epistemic virtues. If in need of an account as to what happened someday on some stage, theater scholars usually give preference to textual accounts.<sup>1</sup> [1]

Not by chance then, theater photography has received rather limited scholarly attention. The illustrative use prevails and is complemented by a theoretical discussion of recordings and their epistemic shortcomings in general.<sup>2</sup> But the volume of such images, published in newspapers, journals, books and on websites, and even in unpublished archives, is not just quantitatively impressive. They also have a significant qualitative influence on the public appearance of the performing arts in modern media societies, particularly in the second half of the 20<sup>th</sup> century. These vast amounts of images, stored away in the archives of photographers and museums, are a vast treasure that provides unique insight into the visual history of the performing arts, in a way that no other medium can offer. [2]

---

<sup>1</sup> Despite a strong historiographical tradition and the location of theater cultures in media environments (cf. Zarrilli et al. 2010), images have rarely been discussed in regard to their problems and potentials as historical sources (cf. Balme 1997).

<sup>2</sup> New research in regard to dance photography (cf. Wortelkamp 2022) or the photography of performance (cf. Dogramaci 2018) has recently changed a similar situation in neighboring fields.

Photography, sociologist Pierre Bourdieu has argued, constructs sociograms, portraits of roles and relations of those performing themselves in front of the camera, their norms and values, negotiating on the fly what is worthy of representation and what not.<sup>3</sup> An image of a family reunion or a village fair is not much different from a theater photograph in that respect. Here too, we come across the self-imaging of a group. Seeing Othello or Desdemona, whether on stage or in a photo, we see not only how that role is staged but also who was hired to play the character. In theater photography we not only encounter an aesthetic reflection of society, but also a part of that society. We encounter a cultural history of the body in these image collections. [3]

But due to constraints of access, which result from the limitations imposed by several features including the epistemic order of existing aids to find and locate entries, these archives have remained rather ›dark‹.<sup>4</sup> Since there is no way to query the images themselves, a scholar who might be investigating the uses of blackface in German theater, to take but one example, must first think of plays like *Othello*, where they seem likely, only then to browse through the productions of such plays.<sup>5</sup> It is the canonical repertory order that governs the search, and what lies outside of it will inevitably be missed since the holdings are just too large and too unwieldy to manually go through in any reasonable amount of time. [4]

New methods of visual search dependent on recent developments in computer vision and machine learning, unimaginable only a few years ago, are about to change this.<sup>6</sup> They allow researchers to avoid the bottleneck of dramatic metadata and search inside the image. They also enable the large-scale processing of such corpora and their study at scale. This is where the epistemic potentials lie. Not only in finding the performative needle in the photographic haystack, but rather in analyzing the haystack itself: its style and content, norms and conventions, clusters and outliers, gradual differentiations. This shifts attention from the relations between individual images and the performances they depict to the visual language of theater photography and how it mediates performance. Studying the media of performance, but as data, adopting 21<sup>st</sup> century technologies to make sense of 20<sup>th</sup> century ones, not only opens up a long-dormant corpus of sources, but also novel perspectives on the subject of the discipline. [5]

Such a distant investigation of theater photography connects to a wider discussion of digital historiography and digital heritage in the performing arts.<sup>7</sup> In particular, it is in line with the systematic development of Digital Humanities tools<sup>8</sup> and the exploration of computational methods<sup>9</sup> in Theater Studies. Methodologically, it can build on role models in Computational Literary Studies<sup>10</sup> and Digital Art History.<sup>11</sup> Technically it relies on machine learning models that are able to recognize likeness and similarity in images by measuring the proximity of vectors in a high-dimensional matrix.<sup>12</sup> [6]

The key question then is, what features should constitute these vectors? In other words, what counts in these images, in the double sense of the word, from a Theater Studies point of view? How do we operationalize the similarity of performances? The answer proposed below is that it's the bodies that count, which in turn raises questions about the epistemological stakes that come with using subjective industrial technologies for human differentiation in the course of Humanities research. [7]

<sup>3</sup> Cf. Bourdieu 1983.

<sup>4</sup> Cf. Jaillant 2022.

<sup>5</sup> According to anecdotal evidence from the head of the FID Performing Arts, a specialized information service and search engine for the performance arts in Germany, seated at the university library in Frankfurt and funded by the Deutsche Forschungsgemeinschaft (German Research Foundation, DFG), this was the most frequent query in 2022.

<sup>6</sup> Related approaches by GLAM institutions to introduce forms of visual search and semi-automatic annotation of digital photography collections build on the same technologies (cf. Resig 2014; Seguin 2018; Lincoln et al. 2020; Han et al. 2022).

<sup>7</sup> Cf. Bay-Cheng 2016; Otto 2023.

<sup>8</sup> Cf. Bardiot 2020.

<sup>9</sup> Cf. Escobar Varela 2021.

<sup>10</sup> Cf. the journal *Computational Literary Studies*, published since 2022.

<sup>11</sup> Cf. the *International Journal for Digital Art History*, published since 2015.

<sup>12</sup> Cf. Kelleher / Tierney 2018.

Undertaken as a feasibility study, this paper intertwines an exploration of theater photography as an epistemic object with a discussion on the methodological challenges and potentials of computational methods in the field of Theater Studies. It aims to lay the groundwork for larger scale projects by showing how such novel methods can make neglected corpora of sources accessible, open up radically new lines of questioning, and stir up dominating disciplinary paradigms. [8]

It is based on an ongoing study of the historical practice of theater photography, conducted in close cooperation with archival institutions, in particular the [Deutsches Theatermuseum](#) (German Theater Museum) in Munich and its director, Dr. Dorothea Volz, as well as contemporary photographers, namely Iko Freese and the agency DRAMA in Berlin. The focus, elaborated here, has been on data survey and computational exploration of representative photographic datasets. This has been conducted using object detection and computer vision libraries like [OpenCV](#) or [Detectron2](#) in Python, making extensive use of [GitHub Copilot](#) on a local environment, and [Google Colab](#) to access GPU capacities and to avoid dependency issues. [9]

Technically, the central challenge that emerges is in the area of pose estimation, in moving from detecting the key points of joints in a two-dimensional image to a three-dimensional model of human bodies that allows for a meaningful comparison. Here, the project builds on the work and advice of Stefanie Schneider<sup>13</sup> in Digital Art History and recent research in human mesh recognition,<sup>14</sup> including the *Skinned Multi-Person Linear Model (SMPL)*<sup>15</sup> and *SMPL eXpressive (SMPLX)*<sup>16</sup> models of the human body, initially developed at the MPI for Intelligent Systems in Tübingen. [10]

Section 2 introduces the scene shot as an object of study and the research agendas traditionally connected with it. Subsequently, Section 3 discusses the theoretical stakes of a distant study of these artifacts from a Theater Studies point of view. This is followed, in Section 4, by a discussion of the operationalization of the analysis of body images. Section 5 addresses the technical challenges that arise from the deployment of industrial *deep learning* frameworks in realizing such an approach and proposes to understand such digital hermeneutics as much as an interrogation of the object as of the machinic gaze. Section 6 concludes by reflecting on the use of computational methods for the discipline of Theater Studies, and the reciprocal contribution of Theater Studies to the transdisciplinary field of Digital Humanities. [11]

## 2. Mediating Performance. The Case for Theater Photography

Theater photographs come in many forms and may show very different things: the private portrait of an admired actress as a cabinet print; a postcard presenting a stately building in the city center; or a large promotional poster showing a striking set design – usually, though, these are not the formats that are considered when the medium is discussed, but rather images of dramatic stagings, made primarily for the press and most commonly used in the context of reviews.<sup>17</sup> [12]

Such *scene shots* of theater in action date back to the early 20<sup>th</sup> century but only became standard practice in the 1950s.<sup>18</sup> They require not only the appropriate light-sensitive film materials to work under stage conditions, but also imply a demand for such images – initiated by regular publication in newspapers, magazines and trade journals. Theater photography thus emerges as a historic cultural practice. It becomes [13]

<sup>13</sup> Cf. Schneider / Vollmer 2024.

<sup>14</sup> Cf. Stathopoulos et al. 2024.

<sup>15</sup> Cf. Loper et al. 2015.

<sup>16</sup> Cf. Pavlakos et al. 2019.

<sup>17</sup> Accordingly, the term ›theater photograph‹ has no defined meaning and is used deliberately to refer to any photography that is discussed in relation to the art of theater – for an introduction into the topic cf. Anderson 2019.

<sup>18</sup> Cf. Meyer-Plantureux 1992.

institutionalized in financial budgets, contractual agreements, educational textbooks, and most notably in designated photo rehearsals, during which actors and directors must endure the clicking of shutters and sometimes even the presence of photographers on stage. Subsequently, the products of this practice, developed overnight by hand, selected and approved the following day by directors and dramaturgs, journalists and editors, enter the news cycle and get distributed – but only after being combined with a caption, a title and some text. In this way, theater makes a visual or rather a multimodal impression on a wider public: on those who couldn't make it to the show because of time, or because they didn't have money for tickets, or because they were disinterested, were living too far away, or were simply born too late. Once the run is past, these scene shots end up in archives and constitute a visual trace that is reproduced and reinterpreted by history and schoolbooks to form part of the cultural memory of theater.

Obviously, such a scene shot is not the performance. It does not even come close to what a live performance is all about: being there together and experiencing the passing of time, the heightened perception, and the plays of meaning that unfold with being addressed from the stage.<sup>19</sup> The photograph, in contrast, is detached from the event, the time and place where it happened. It condenses time to a single moment and projects space onto a plane. It comes without sound, speech, and often lacks color. Above all, it has its very own aesthetics.<sup>20</sup> It is something completely different – and yet it is not. Traditionally, it misses the autonomy of a work of art like a painting, that is admired and valued in and of itself, independent of the object depicted. Regardless of its aesthetic qualities, the scene shot remains dependent on the performance it portrays and whose history it constructs. Even though there are theater photographers of considerable renown with distinctive visual language, who have had their works presented in solo exhibitions and illustrated volumes, this work remains inextricably linked to the fame of the theaters and directors who were photographed.<sup>21</sup> Even the most prominent and confident photographers unanimously describe their artistic practice as translating a performance that retains priority, and whose integrity they aim to preserve. They clearly understand their work as an *applied* form of art, that comes with only a limited claim to autonomy.<sup>22</sup> [14]

Unsurprisingly, the principal interest in these images comes from the wider field of theater, including the academic disciplines of Theater, Performance and Dance Studies. The dominant use case tends to be representation: scene shots picturing what is playing and what has been playing, most often in an illustrative manner, as a supplement to a text without further comment.<sup>23</sup> A relatively small set of photos is used to reproduce a largely canonical repertoire, while the majority of images are put away in the archives. This is generally overshadowed by a dominant theoretical discussion about the axiomatic lack of epistemic validity of these sources, which is based on an essentialist distinction between the ontologies of live art and media culture, that, in turn, is grounded in a modernist aesthetic framework which lies at the foundation of the discipline.<sup>24</sup> It is only fairly recently that research on documentation,<sup>25</sup> in particular of performance art,<sup>26</sup> has demonstrated the crucial role that recordings play in the emergence and establishment of the works and the institutions of the performing arts. [15]

In principle then, scene shots are of interest less as aesthetic objects, but as epistemic things:<sup>27</sup> they are a way of knowing about a performance and making sense of theater, basically resembling the samples a scientist brings back from her lab work or field trip. Their primary objective is to *represent* what is out there, and through this representation make accessible and comprehensible what would otherwise be too unwieldy to [16]

<sup>19</sup> Cf. Fischer-Lichte 2004.

<sup>20</sup> Cf. Diekmann 2010.

<sup>21</sup> To name only one example: the work of photographer Ruth Walz is inextricably linked to the work of the Schaubühne under Peter Stein.

<sup>22</sup> This observation relies on a yet unpublished series of qualitative interviews with contemporary theater photographers.

<sup>23</sup> Cf. standard theater history overviews like Brauneck 2007.

<sup>24</sup> This has been an ongoing discussion with the initial antagonistic positions represented by Phelan 1993 and Auslander 1999.

<sup>25</sup> Cf. Reason 2006.

<sup>26</sup> Cf. Gusman 2019.

<sup>27</sup> Cf. Rheinberger 2006.

deal with as a whole: the performance – that, in so far as it is not an artifact like a painting, a sculpture or a movie, shares the fuzzy ontological boundaries of a jungle, a galaxy or an organization; matters dealt with by the sciences.<sup>28</sup>

As a proxy of a performance, the scene shot shares none of its aesthetic qualities but gives it a public presence, i.e., a mediated reality beyond the here and now to an otherwise spatially and temporally restricted event – and in doing so, it articulates at the same time a more abstract idea of the institution that produces such performances, the theater, and what it might be all about. Joining in on a centuries-old tradition of announcing and reporting theatrical events by means of writing and drawing, the scene shot is not without priors and is not unaccompanied. But it gives a new twist to the mediation of performance, due to its specific mediality, i.e. its indexicality; stressing the durational qualities of the performing arts by fixing the moment and offering the public a detailed visual impression of the stage action, that allows for new ways of valuing, judging, studying, or more generally speaking, knowing about and relating, to the performing arts and what matters about them.<sup>29</sup> [17]

Here, too, representing means intervening,<sup>30</sup> knowing takes part in the making. The camera turns the fleeting moment that made sense only within the frame of performance, a scene, a play, a show, into an image of some body in action, to be taken outside the situation and relocated into showcases, newspapers, or social media. It becomes an icon, a sign for a performance, a company, an institution or culture in general; a sign though, that still bears witness, and might inhabit what Roland Barthes called the *punctum*, something that escapes the staging efforts of actors, directors and the photographer and does not fit into the semiotic order and strikes the spectator in uncanny ways.<sup>31</sup> [18]

The key question regarding theater photography, then, is a question of *mediation* in the broader sense of Actor-Network theory:<sup>32</sup> how does the shot, its production, distribution, and reception contribute to the articulation of theater in a specific historical situation? I.e., how do these images, or rather the practices of *imaging*, of image-making, give form, presence and meaning to particular performances beyond the social interactions of the in-situ event that their aesthetics are grounded in? How do media make theater matter? How does the emerging visual knowledge contribute to constituting theater as a wider social institution and epistemic community? And how do performances become effective beyond the here and now of performance due to their visual presence in changing media environments? [19]

Here, theater photography is of interest not as an artifact, but as a cultural practice: as a shared and embodied way of seeing and doing things, held together by a set of largely implicit values, beliefs and conventions that organize the practice and its products; centering the frame on the interaction of actors, pressing the button the moment that faces are visible but mouths are shut; choosing images that reflect the cast of the show; adding captions that give names to both actors and characters, etc. [20]

Consequently, attention shifts from any (exceptional) single image and its relation to a corresponding (canonical) performance, which has long been at the center of academic attention, to the relations between the images themselves: the multitude and its order. What do these images *look* like, i.e. what visual structures [21]

<sup>28</sup> For an understanding of how science works cf. the field of Science and Technology Studies, in particular the Actor-Network theory branch, most prominently in Latour 1999. Early ideas on how to apply these ideas to the field of Theater Studies are in Otto 2020.

<sup>29</sup> Video recordings in contrast have had a much smaller influence on the public perception of theater, due to the lack of distribution, despite offering a much more information-rich impression of a performance.

<sup>30</sup> Cf. Hacking 1983.

<sup>31</sup> Cf. Barthes 1985.

<sup>32</sup> Cf. Latour 2005.

do they have in common? What do they *show*, i.e. what kind of entities do they contain? And what do they *tell*, i.e. how do they give meaning to what they depict? Moreover, how does all of this differ in relation to place, time, and genre?

### 3. Surface Matters. Distant Viewing for Performing Arts

One way to understand the mediation of performing arts is to observe closely how it works in daily practice. [22] A different approach, the one pursued here, is to study the artifacts it produces: not a single image of an exceptional production, but the hundreds of thousands of images in the archives of photographers, most of which are private, but some historical ones can also be found in museum collections.

Due to their commercial nature, these collections are ideally suitable for a study at scale. They are usually [23] well-organized and documented, exhaustive, and share relatively uniform standards regarding data as well as metadata.<sup>33</sup> This allows for largely automated processing and for an easy merging of collections from different photographers and times, the only challenge being the sheer size of these inventories. The 63 photographic estates held by the Deutsches Theatermuseum in Munich, for example, range from a few thousand to several hundred thousand negatives each, amounting to several million items in total. The archive of a contemporary photographer like Iko Freese, who started working in the 1990s, comprise millions of negatives and even more digital files. No PhD student would be able to view such quantities in their lifetime, let alone during a funding term; and, even if by a clever strategy one was able to pick a meaningful sample out of the pile, a comparative study by hand and eye would be out of the question without multiple assessors. Only by digital means, i.e. through numeric representation, computational processing and statistical evaluation, does the scale of this corpus become manageable, and thereby epistemically valuable.

Such studies of scale, that pursue a statistical or more generally computational analysis of large corpora, [24] are often referred to as *distant* – in contrast to the proximity of the hermeneutic encounter with a unique artifact.<sup>34</sup> Initially such methods were developed in relation to textual artifacts, which have the advantage of consisting of discrete units that are relatively easy to standardize. Building on research in computer linguistics and using pipelines for Natural Language Processing, the field of Computational Literary Studies has elaborated a plethora of approaches like digital stylometry or topic modeling that basically start with counting words. Methodologically, such approaches rely on assembling a representative corpus, cleaning the data in what is often the most time-consuming step, and then analyzing statistical distributions and correlations in that data. This is followed by the visualization, contextualization and interpretation of the results.<sup>35</sup> Central to the approach is the question of what relations to measure in the data, which boils down to translating qualitative concepts from the Humanities into some kind of quantitative variable; a process that Franco Moretti, picking up a term from theory of science, has designated as ›operationalization‹: »building a bridge from concepts to measurement«. <sup>36</sup> But since such a translation is not without significant gains and losses, it is both potentially problematic and epistemically productive. For Moretti, who originally coined the term ›distant readings‹, though in a slightly different context, the approach accordingly comes down to an epistemic intervention against the ›secularized theology‹ of hermeneutics<sup>37</sup> and is associated with three trajectories: a move outside of the canon, towards serial productions and towards long-term trends.<sup>38</sup> And it is in this respect in particular that the umbrella term ›distant reading‹ is still employed and helpful to describe very different computational approaches.<sup>39</sup>

---

<sup>33</sup> The EXIF metadata includes information about camera models, technical parameter, as well as date and time.

<sup>34</sup> Cf. Moretti 2016.

<sup>35</sup> Cf. Drucker 2014.

<sup>36</sup> Moretti 2013.

<sup>37</sup> Cf. Moretti 2000a, S. 208.

<sup>38</sup> Cf. Moretti 2000b.

<sup>39</sup> Cf. the COST Action project [Distant Reading for European Literary History](#).

More recently, with advances in computer vision, image data has come within reach of the Digital Humanities and analogous approaches of ›distant viewing‹<sup>40</sup> have been proposed in regard to (audio-)visual media. In particular, the fields of Digital Art History, focused on classical art works, and of Cultural Analytics,<sup>41</sup> studying popular images with a tilt toward Cultural Studies, have begun to analyze images at scale. But images, even when digitally represented as a three-layered pixel matrix, are not semantically discrete, and segmentation becomes a problem: it is not obvious what to count, and the corpus is ambiguous, since it is not always clear where and between what to draw the line, in particular when it comes to art works. If literary style is translated into the words used, pictorial style is much harder to pin down and a much more malleable category, especially if the term is not reduced to painting. That is why the question of features and their detection becomes pivotal when computing images: these need either to be ›engineered‹, i.e. algorithmically defined by humans, or acquired through a training process of a machine learning system. In consequence, approaches of ›distant viewing‹ in contrast to ›distant reading‹ are much more concerned with questions of representation. [25]

When it comes to performances though, a distant approach of inquiry gets even trickier. Working with texts and images, both Computational Literary Studies and Digital Art History work on the artifacts their disciplines are built on, the central challenge being how to make them computable in a meaningful manner. But a performance cannot be digitized like a text or an image; and many scholars will still tell you that such an image doesn't really count in the first place.<sup>42</sup> Dealing with works that are events, the closely related disciplines of Theater Studies, Performance Studies and Dance Studies have an immaterial basis and are reliant on a variety of left-overs that are an ontological mess, from the perspective of aesthetic theory and data science alike. From review snippets, photo albums, costume parts via director's books, actors' diaries, set models, to amateur films and television productions, it's a plethora of different kinds of objects, from all kinds of different worlds, impossible to squeeze into a unified metadata model, all of them inevitably radically incomplete and fragmentary, the more so when put together, and none of them with a claim to epistemic primacy. Even a multi-channel video recording of a performance, despite being more captivating, is anything but an accurate representation and rather a thing of its own, due to elaborate camera work, editing techniques, and sound mixing. There are strong reasons to doubt whether the higher density of information contained in such a recording, and any re-construction based on it, would lead to a better understanding of the significance of that performance compared with the black and white scene shot and a review in a local paper. [26]

Accordingly, from a Theater Studies point of view, any study of images is in principle already some kind of distant reading and must expect objections as to its epistemic adequacy.<sup>43</sup> If the study of these media is already a secondhand investigation, its digital processing at scale will be even more so: both are in danger of missing, or, worse, distorting, what actually matters about a performance, the ephemeral qualities of the lived experience. No need for pictures, or counting pictures, if all that counts is performance – and if performance is precisely what eludes (ac)counting and (ac)countability. As long as these images are understood only in relation to what they have to say about a temporally and ontologically prior performance, the results of any image analysis necessarily remain superficial. [27]

That is to say, these objections have their own epistemic prerequisites. They are grounded in a modernist paradigm of Theater Studies that is built on the idea that the discipline is defined by the nature of its subject matter, which in turn is understood as a specific kind of art, defined by its (im-)materiality. Performance is devised in opposition to the text, the artifact, but still as a work of art, i.e. as a somehow self-contained entity, [28]

<sup>40</sup> Cf. Arnold / Tilton 2019.

<sup>41</sup> Cf. Manovich 2020.

<sup>42</sup> Despite novel approaches and a lively methodological debate (cf. Balme et al. 2020; Wihstutz / Hoesch 2020), this paradigm is still dominant, as experienced recently by me, when a funding application was categorically rejected by a reviewer on the grounds that any such distant approach necessarily misses the subject of the discipline.

<sup>43</sup> Consequently, most Digital Humanities initiatives and projects in Theater Studies are focused either on the collection and modelling of performance metadata or the visualization of performance spaces.

that can be analyzed as such, so that it gets theoretically contained in and constrained by the here and now of an event, methodically cut off from medial and social entanglements. To stay within the hermeneutic tradition, Theater Studies needs to construct performance as a quasi-artifact, an ephemeral but at the same time enclosed, complete and stable *thing*.

A distant study of theater is indefensible within this paradigm, because it contributes little to the reading of performances as artistic works – and has to seek its fortune outside of it. It is less filling a research gap than offering a novel outlook by, at least heuristically, turning its back to hermeneutics and embracing empirics, tentatively moving from Theater Studies to Theater Science. Performance returns as a problem; more in need of explanation than of understanding. Instead of wondering what is *theoretically* lost in the mediation of performance, it examines what is *practically* gained: how do media overlay, organize, and generate the matter, meaning, and agency of individual performances and consequently theater as an institution?<sup>44</sup> How does performance happen beyond the assembly of people in the here and now? And how does mediation make theater work in the first place?

[29]

## 4. Body Counts. Operationalizing the Theatrical Gaze

Such an empirical stance implies taking media seriously and taking a fresh look to see what is actually in the media archive, instead of subordinating its contents to the repertory order and reducing them to the representation of a literary prior. It requires a break with the emic logic of the field, that is shared by venues, journals, archives, and often academics, in order to adopt an etic stance, an outside perspective, that insists in the tradition of Theater Studies on the primacy of performance.

[30]

Starting with a data survey, a closer look, that adopts the engineer’s point of view, helps to establish the necessary distance: what are the features to be identified, the values to be measured, the relations to be analyzed? Operationalization implies formalizing and making explicit the necessarily fuzzy and often implicit categories of the Humanities.<sup>45</sup> And already this process of making computable what is often understood to be incalculable, a translation that is inevitably a transformation, holds potential for insight.

[31]



Figure 1: Tryout of Detectron2 / OpenCV object detection (left), 2D pose estimation (middle), and image segmentation (right). [Production: *Die Weber*, Volksbühne Berlin 1993, director: Frank Castorf, photographs © Iko Freese]

<sup>44</sup> Cf. Balme / Fisher 2020.

<sup>45</sup> Cf. Impett / Moretti 2017.

What stands out from this perspective is, first of all, that it is not plays we encounter in these images but bodies in action (cf. Fig. 1). It is the human body that dominates image compositions. Not unlike film shots, these images are visually centered on the human figure in a highly conventionalized mode of composition and can be categorized according to the number of bodies contained or the framing and cropping of the body sections depicted (long and medium shots, closeups, etc.). Almost all pictures are taken from the front, averted faces or blurred gestures are rare, a cropped figure is unlikely and the absence of humans an utmost exception. Heightened expressions, pronounced postures, and theatrical costumes draw attention and carry particular meanings. In particular, the relation of these bodies to each other constitutes the scenic content by mapping social situations (cf. figs. 2).

[32]

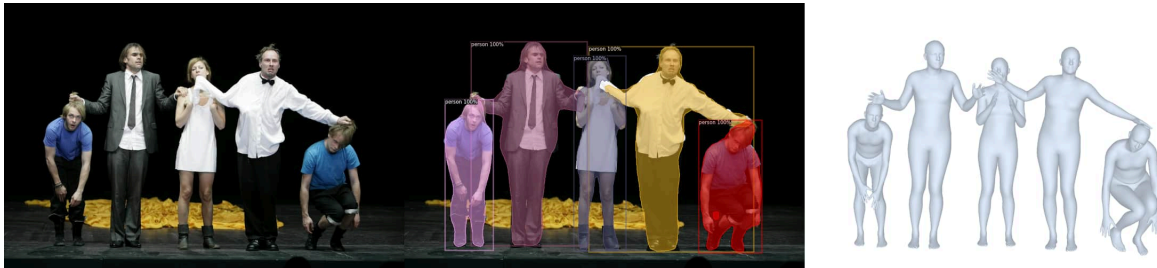


Figure 2: Example of scenic constellation in plain photograph (left), annotated by object detection (middle), and reduced by mesh recognition (right). [Algorithm ScoreHMR, production: *Anatomie Titus Fall of Rome*, Deutsches Theater 2007, director: Dimiter Gotscheff, photographs © Iko Freese]

As long as the bodies are our contemporaries, we perceive these images to be simply an extension of the theatrical present. But the more time passes, the more distant they get, the more they turn into the »archaeological mannequins« that Siegfried Kracauer encountered in early photography.<sup>46</sup> That is because these scene shots are basically press photos that only make sense in relation to an event being depicted, not as an autonomous work of art, an event in itself. Not unlike sports photos they have a denotative meaning that is hard to understand without knowing the rules of the game, the line-up of players, and the specific match. But they can also acquire a larger, connotative meaning regarding not only a certain performance, but also the historical moment in which it took place. That is because they depict not only characters but also actresses, not only artistic figures but also social bodies belonging to history proper, showing not only an aesthetic event but a historical one at the same time. The latter comes to the fore the more time passes and the more alien in style the bodies become.

[33]

From a theatrical point of view, it is less the image, then, but the body, or more precisely the *historical bodies in action* at the intersection of aesthetic and cultural history, that emerge as the central unit of interest, whose computational description is the central challenge and variable of a distant viewing of theater photographs – both in relation to the repertory order, but also independent of it. How do they relate, how do they pose, what do they look like and how many are there?

[34]

Such an analysis could be systematically divided into the following aspects:

[35]

- A scene shot first of all contains a certain number of bodies – a very simple parameter that can help to differentiate between soliloquies, constellations, and choirs.
- Each of these bodies can be characterized by several distinct but related features: (a) physical characteristics associated with structures of bones and tissue, plus descriptors like thin or tall, including attributions of social categories like gender; (b) body poses, hand gestures and facial expressions can be described in terms of joints and muscles, as well as descriptors and interpretations like sentiments

<sup>46</sup> Cf. Kracauer 1977.

associated with such poses, gestures, and expressions; (c) the cut, material and color of clothes and hair being worn as well as the associated style.

- Taken together, the bodies are located in relation to each other but also towards things, having a spatial position and orientations of different body parts that create relationships in terms of proximity, gaze, and touch, and translate into descriptions such as ›separate‹, ›facing‹, ›averted‹, etc. – is the body standing alone or in contact, physically interacting with things and with what things, directed toward whom, looking where?
- Finally, there is the photographic depiction associated with camera parameters like position, angle, lens, focus, framing, plane, overlay, lighting, that produces an impression of the body, which again is in need of a qualitative description for an actual understanding.<sup>47</sup> This would allow for a numerical description of scenic content contained in the image, independent of the dramatic context.

On a second level, these bodies do not only make sense in relation to others within the image, in the intra-pictorial frame, but also in an inter-pictorial context, in relation to other, more or less similar bodies, that come before and after them. On the one hand there is the series of production photographs that any particular body appears in, ordered by shot number and, more recently, datetime entries (cf. Fig. 3). This maps onto the continuous but intangible timeline of a staging, possibly reflected in the timecode of a video recording and the structure of a playbook, in the case of dramatic theater eventually mapping to a play, an act, a scene, a line. On the other hand, this intradiegetic dramatic order is embedded within the repertory order on the institutional level, the same play being enacted again and again in changing times, by different people, in ever new variations, so that each theatrical body is related to other bodies that appeared as the same character, as well as to other characters represented by the same body. It is at the intersections of these series, one defined by the dramatic order of a specific production, the other by the repertory order of the multitude of productions, that these bodies make sense. And a simple visualization of either of these, giving either a visual footprint of a production or displaying the scope of a character's interpretation, provides a helpful tool (cf. Fig. 4). Further options arise from the possibilities to detach groups of bodies from sets of images and visually re-organize them according to the dramatic and repertory metadata.

[36]

---

<sup>47</sup> Here it is of the utmost importance to distinguish between characteristics that are to a great extent syntactic and can be measured, like length of arm in relation to length of spine; and those that are cultural and have a semantic overlay, despite being based on such measurements.

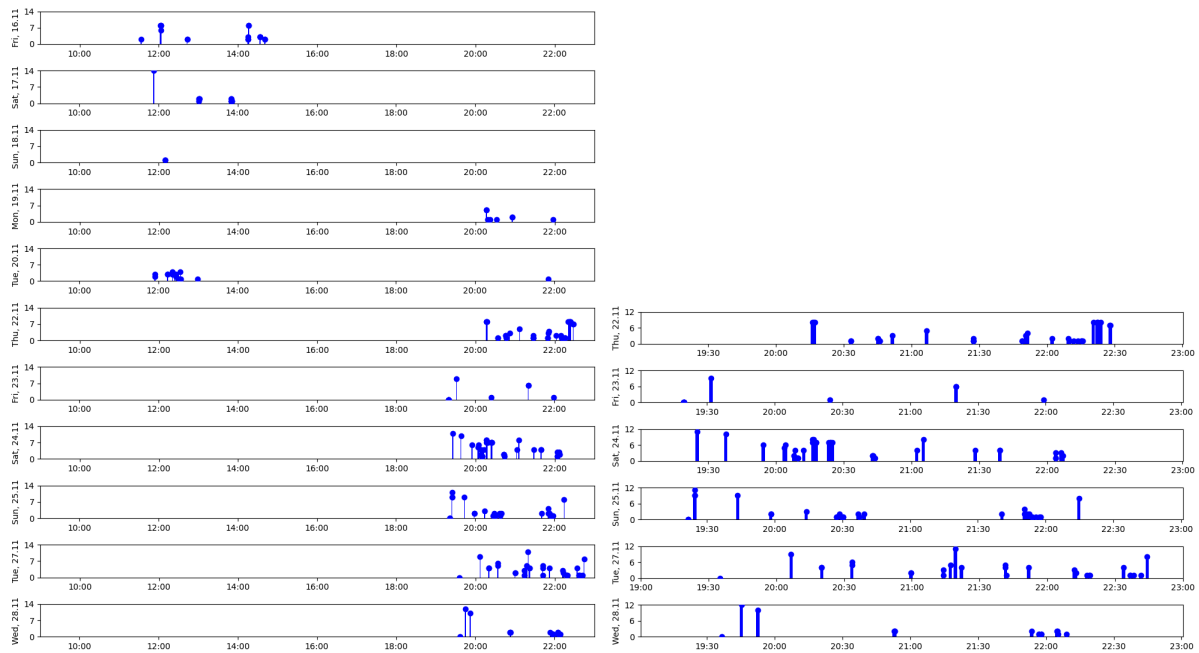


Figure 3: Bar charts of the shot times in the sample set of photos relating to *Anatomie Titus Fall of Rome* in photographer Iko Freese's private archive. This conveys not only an idea of the photographic practice (left), but also shows what moments in the play stick out and are repeatedly pictured (right). [EXIF metadata extracted and visualized with Python: Ulf Otto 2026]



Figure 4: Visualization of all photos taken of the *Anatomie Titus* production in order of time taken (left), reduced to body meshes (right). [Human mesh extraction with help of ScoreHMR: Ulf Otto 2026]

Then again, the crucial potential of the computational approach lies in scale, and scale implies to take these bodies, at least heuristically, out of context, i.e. out of the dramatic and repertory order, to move beyond the frame of a specific play or production. Provisionally detaching the bodies from their scenic meaning allows us to zoom out instead, to examine what kinds of bodies, what norms and formulas rule their expression, into what groups they are clustered. What characteristics distinguish these groups, what outliers do they

produce and how do they change over time? What measures can be developed to compare forms of bodily expression? What insights does a visual analysis of 20<sup>th</sup> century acting styles provide? How does an inventory of poses or typology of scenic constellations look like?

Though this line of questioning might eventually relate back to the understanding of some canonical productions, it initially and essentially leads to insights about theatrical practice and its medial portrayal in its breadth and gradual development. Taking these bodies out of their aesthetic context emphasizes the social and historical place of these bodies and relocates aesthetic questions within a social frame. [38]

## 5. Bias at Work. Deploying Machine Learning for Performance

If what is of interest from a Theater Studies point of view is not the image itself, but what is in it, what it shows about the show, namely the bodies, or more precisely, the bodies in action, then the task at hand is, speaking in technical terms, a matter of *information extraction*. [39]

Given the millions of artifacts that a distant study is concerned with, the ›too-much information for any PhD student to process in a lifetime‹, the central question is how to reduce these artifacts to a simpler, lower-dimensional representation, which retains those features that are meaningful for the questions at hand, to be able to compare them. [40]

Early computer vision systems allowed for a distillation of basic image feature like edges by matrix operations, but advances in machine learning have enabled important new functionalities: ›object detection‹ allows us to find human bodies as well as a limited repertoire of things in the image, while ›instance segmentation‹ can crop these entities; and ›image classification‹ identifies certain classes of images. Combined with basic information about the width and height of a frame this allows for an approximate categorization of the image. Furthermore, face recognition facilitates the ability to find actors in the images and, combined with the cast list, allows one to attribute character names. Most importantly ›pose estimation‹ provides key points for the joints of the skeleton. And while 2D key points are only partially suitable for comparison of poses, they help in deriving complete 3D body models by what is called ›human mesh recognition‹ and allow us to define metrics for body expressivity. Though all of these systems struggle with occlusion, and it is advisable to have a human in the loop to filter results, they work well and fast with photographic data and allow for rapid (semi)automatic processing. [41]

Related systems add semantic information: the Google Cloud Vision API, as one example, detects objects and faces, but also attributes ›sentiments‹ to the latter, ranging from ›joy‹, ›sorrow‹, ›anger‹ and ›surprise‹ to ›exposed‹, ›blurred‹ and ›headwear‹. It further assigns two kinds of labels to the image as a whole: one marking the things or events possibly depicted, the other rating the image in regard to the problematic rating of the content as ›adult‹, ›spoof‹, ›medical‹, ›violence‹, ›racy‹. [42]

All of these rely to some degree on multi-layered neural networks, so-called *deep learning*, that has developed unexpected capacities in dealing with fuzziness by scaling the size of networks and training data, as well as developing new methods of training. One basic technology, *Convolutional Neural Networks (CNNs)*, has stacked layers that represent different levels of image abstraction: low-level features marking edges, mid-level features reflecting shapes, and high-level features being associated with objects. Basically, these systems abstract common patterns from the data given to them but lack a world model. An early system, trained on hundreds of images to recognize birds, was very successful in spotting birds, but, like Grover from *Sesame Street*, also mistook planes, rockets and superheroes for birds, because it recognized a bird as a patch of dark on a large area of blue. Likewise, giving a set of historical theater photographs to an open-source model like Detectron2 from Meta finds a lot of baseball bats in the hands of kings; a generative model, prompted for a [43]

depiction of a theater performance, produces images all associated with high school musicals. Described as »stochastic parrots«, as a famous paper in the field of Critical AI put it,<sup>48</sup> these models represent the »optical unconscious« (Walter Benjamin) of the image set they were trained on.

Inevitably, and necessarily so, the machine gaze is biased. Inevitably, because a machine learning model is generally a socio-technical system that is built from human work and human data. Therefore, the machine gaze is as much a social gaze as the human gaze and can be no less biased. Moreover, bias is exactly what makes the machine gaze work: learning to recognize a bird as a bird, a baseball bat as a baseball bat, comes down to acquiring a cognitive bias toward that conclusion. [44]

This gets messy when human bodies and their categorization on grounds of visual characteristics are at stake, as is the case in dealing with theater photography. The risk is that racist, misogynist, ableist biases, contained with high probability in training data sets, will get reproduced and reinstated, projected onto the materials to be analyzed. Let's say we want to investigate how constructions of femininity have changed on stage during a certain period. We compile a data set with female characters for that purpose, but the model we use has a very limited notion of what counts as feminine, discarding from the very beginning the figure we might be most interested in. Not only might the training data have provided it with a culturally very specific idea of femininity, but it is also already built on a binary logic of distinction, that is itself historical (cf. Fig. 5). This applies not only to semantic distinctions, between, say, a »smile« or a »frown«, »kneeling« or »squatting«, but also more »technical« descriptors, that provide the positions and rotations of the joints and give a numerical and continuous characterization of build and pose, that might appear more »neutral«. But even these models, that are used to represent and recognize bodies, have been abstracted from a set of given bodies that, in the case of the prominent SMPL-model for example, include fat and thin, tall and small people, but neither children nor certain genetic variations, inevitably representing a limited range of variation of human bodies that is not without exclusions. If a system is trained on social data, it inevitably contains bias. [45]



Figure 5: Misgendering is a common issue when dealing with theater data, which can be taken as a starting point for an exploration of gender construction in 20<sup>th</sup> century performing arts. [Screenshot of Detectron2 with OpenCV object detection on video recording of *The Persians*, Salzburger Festspiele 2018, director: Uli Rasche]

<sup>48</sup> Cf. Bender et al. 2021.

In this respect the output of the machine is generally no different from the ›data‹ a phenomenologist brings back from the attendance of a performance, or an ethnographer from a field study: all are but ›capta‹,<sup>49</sup> empirical evidence, some kind of ›measurement‹ made by instruments, even if that instrument is the researcher's body; and such measurements neither speak for themselves, nor do they show the world as it is, but are as much in need of an in-depth discussion of provenance and significance as any poem is. Neopositivist dreams of an isomorphic mapping of reality by apparently neutral machines are misguided, because it is exactly with its denial, in the claim of an untouchable God's eye perspective, that bias gets epistemologically problematic.<sup>50</sup> [46]

In this context, digital hermeneutics are not that different from the classical ones, they enact an epistemic interaction, an encounter of a gaze in the present with an object from the past – only that the gaze does not live in the body of a scholar, but in the weights of a neural network. Both are biased, only in different ways, since both are embodying society. Any insight into the object of inquiry is also an insight into the place of that inquiry. Consequently, the questions addressed to such a system should not be: »How many men are in there?«, but rather: »How many of the beings in there does a specific historical model classify as male with what probability score?« More generally: »What do computers trained on contemporary internet photos see in theater images of the past?« From a Humanities standpoint the use of computer vision inevitably raises the question of the gaze. And just as in case of a ›classical‹ performance analysis, things start to get interesting when the machine gets nervous, stumbles, because things do not fit neatly into the accustomed, i.e. hegemonic, perceptual frames and categories, when even in the absence of occlusion and under best lighting conditions, distinctions get blurry. It is in challenging the observer's gaze, negotiating aesthetic norms, and implied social categories, that the encounter with art gets interesting (cf. fig. 6). [47]

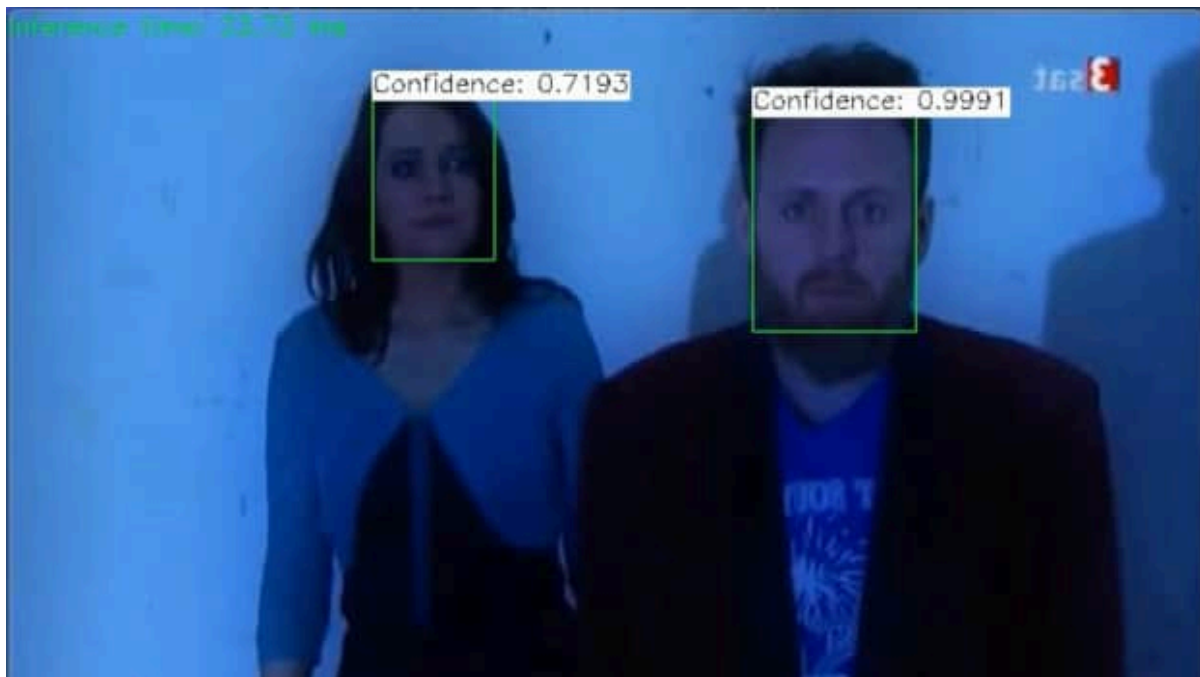


Figure 6: Confidence intervals in face recognition can help to understand staging patterns: here, the thinking male in the front, accompanied by the caring female on the second layer. [Screenshot of Detectron2 object detection on video recording of *Ivanov*, Volksbühne Berlin 2014, director: Dimitar Gotscheff]

<sup>49</sup> Cf. Drucker 2011.

<sup>50</sup> Cf. Haraway 1988.

So, if we need to resort to industry, to what Silicon Valley has to offer for this kind of research – due to the resource intensive development and training of such systems – and have to look through the eyes of companies like Meta, Google, or OpenAI, we import bias. But this does not necessarily equate to an epistemic disadvantage. If instead of trying to fix the systems, we pay close attention to where they fail, the bias of the machine is actually an opportunity to engage with such research approaches:<sup>51</sup> First of all, because it generally readdresses the question of bias in the Humanities. It contrasts and thereby relativizes the embodied social position that, in most cases, is shaped by an upper middle class upbringing and usually still taken tacitly for granted despite all feminist epistemology.<sup>52</sup> It introduces a different kind of bias, manufactured not by parents and teachers but by software engineers and product owners, and which is oriented more toward popular and profane US-mainstream culture, relativizing the academic positions in the front rows. And, by diversifying bias, it invites a different gaze on the object of study that is rarely represented by the humans doing the research; it complicates the picture. [48]

Finally, this helps us to get these systems themselves into view in the Humanities. Employing them as instruments for research implies an investigation of their epistemic workings: confronting a machine learning model with a theater photograph equates to a historical Rorschach test and puts AI on the couch. While IT engineers are putting all their efforts into closing the ›semantic gaps‹ the systems inherit, the Digital Humanities' job is to critically examine how these gaps are bridged. [49]

## 6. Conclusion and Outlook

Based on an examination of the massive but neglected corpus of theater photography, this paper discusses some of the opportunities and challenges of computational methods for Theater Studies. Although it has to be kept in mind that performing arts heritage exists through a heterogeneity of sources, photography offers an ideal starting point because it shifts the focus away from the literary order and reflects the primary visual mediation of the performing arts in the 20<sup>th</sup> century. [50]

Coming relatively late to the Digital Humanities, the discipline can build on methods of distant reading, well developed within the framework of Computational Literary Studies as well as related and more recent approaches in the fields of Cultural Analytics or Digital Art History. Therefore, a parallel field of Computational Theater Studies needs to be located within the decidedly transdisciplinary horizon of Digital Humanities on the one hand. On the other hand, it faces specific disciplinary challenges closely related to its immaterial object of study, performance, that also leads to specific approaches. Turning to big data and empirical methods, statistical evaluation still constitutes an epistemic provocation within the leading disciplinary paradigm, as it foregrounds medial materials whose epistemic status is still generally considered questionable. Accordingly, the potential of zooming out lies outside of this paradigm, in an underexposed set of questions related to the mediation of performance. Mediation here is understood not only in the sense of Media Studies, in the recording, distributing and processing of information by technological means,<sup>53</sup> but also in the sense of making something real, as a pragmatic sociology in the manner that Actor-Network theory understands it.<sup>54</sup> Leaving the proximity of the action, a distant study of performance implies a repositioning of research that is synonymous with a defamiliarization with its object of study. It implies reconsidering performance not as a given whose meaning needs to be deciphered in an individual and intimate encounter with the work of art, but as a societal problem to explain: how are performances made into something that matters for others? [51]

---

<sup>51</sup> Cf. Impett / Offert 2022.

<sup>52</sup> Cf. Harding 1993.

<sup>53</sup> Cf. Kittler 1987.

<sup>54</sup> Cf. Latour 2005.

If making sense of theater through the products of a camera intervenes in the theatrical order, so does employing machines to extract features from these photos. Counting changes what counts, and introducing a form of (ac)counting into an area often celebrated for its un(ac)countability has repercussions for its epistemic order, that despite all modernist, performative and post-dramatic tilts is still essentially the repertory order that subjugates all things theatrical under a performance, whose production (understood as a function of actors, director, texts, place, time) then qualifies as a work of art. [52]

But already the medium of photography, a highly serial form of conventionalized aesthetic patterns, dominated by its journalistic use case, embedded in multimodal journalistic contexts, and being the outcome of an institutionalized professional practice, stands in stark contrast to this logic of the art world. Even more so does its study at scale by help of computational methods. [53]

Studying visual media from a theatrical point of view, then, does not aim at the aesthetic structures of certain modalities like text, sound or images, but at the performative content of such media. In particular, it's the bodies that count; that is, both the aesthetic bodies, at home in theater history, that derive their form and meaning from the artistic decisions made, as well as the congruent social bodies of the actors, that are part of history proper. It's in relation to other bodies within the image that these bodies matter, but also in their inter-pictorial relations that connect bodies from different times and places. [54]

This agenda can draw on large stocks of photographs, organized by a shared archival logic and possessing a relatively standardized aesthetic form, making them ideally comparable. But due to the scale of the materials it has to resort to image processing by computer vision, built on machine learning models provided by big industrial players. The use of such ›stochastic parrots‹ for digital hermeneutics requires a critical methodological reflection, but offers a decentralizing of the Humanities' position, and insights into contemporary digital culture, because subject and method question each other. [55]

At the same time, the research design of such a digital hermeneutics stays within the Humanities frame, in that it is largely explorative in character: the survey, the aggregation, the enrichment, the analysis, the visualization, and the interpretation of the data concerned is not done according to a given recipe to provide a true or false answer to a question fixed beforehand; rather it is an iterative and flexible process, ›agile‹ to put it in software engineering lingo, that proceeds step by step, turns in circles, constantly evaluates itself and slowly unveils the questions hidden within the object of study. [56]

By proving the viability, both theoretically and technologically, of such an approach, addressing real and imagined objections and discussing its methodical challenges, this paper hopes to have laid another brick of the groundwork for a Computational Theater Studies field. But to further pursue such studies, both medium- and long-term initiatives are needed on the level of infrastructure: first, it requires the compilation and enrichment of sufficiently large corpora of theater data, through the digitization of historical materials and the collection and preservation of digital data. Second, a new generation of scholars has to acquire data literacy and statistical skills to enter into productive working relationships with colleagues from DH and IT. Third, scholars need to get accustomed to new forms of collaborative research cultures that have role models in the engineering disciplines. What might also help would be a slight re-orientation of the discipline, from asserting disciplinary distinctiveness in theory and toward making transdisciplinary alliances work practically. [57]

## Bibliography

- Joel Anderson: Theatre Photography Between Theatre and Performance. In: *Focales* 3 (2019). DOI: [10.4000/focales.803](https://doi.org/10.4000/focales.803)
- Taylor Arnold / Lauren Tilton: Distant Viewing: Analyzing Large Visual Corpora. In: *Digital Scholarship in the Humanities* 34 (2019), Supplement 1, pp. 3–16. DOI: [10.1093/llc/fqz013](https://doi.org/10.1093/llc/fqz013)
- Philip Auslander: *Liveness: Performance in a Mediatized Culture*. London 1999. [GVK record]
- Christopher Balme: Interpreting the Pictorial Record: Theatre Iconography and the Referential Dilemma. In: *Theater Research International (TRI)* 22 (1997), no. 3, pp. 190–201. [GVK record]
- Christopher Balme / Tony Fisher (eds.): *Theatre Institutions in Crisis: European Perspectives*. London 2020. [GVK record]
- Christopher Balme / Berenika Szymanski-Düll / Lorenz Aggermann (eds.): *Methoden der Theaterwissenschaft (= Forum Modernes Theater, 56)*. Tübingen 2020. [GVK record]
- Clarisse Bardirot: Theatre Analytics: Developing Software for Theatre Research. In: *Digital Humanities Quarterly* 14 (2020), no. 3. 25.09.2020. HTML. [online]
- Roland Barthes: *Die helle Kammer. Bemerkungen zur Fotografie*. Frankfurt / Main 1985. [GVK record]
- Sarah Bay-Cheng: Digital Historiography and Performance. In: *Theatre Journal* 68 (2016), no. 4, pp. 507–527. [GVK record]
- Emily M. Bender / Timnit Gebru / Angelina McMillan-Major / Shmargaret Shmitchell: On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In: *FACCT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. 2021, pp. 610–623. PDF. DOI: [10.1145/3442188.3445922](https://doi.org/10.1145/3442188.3445922)
- Pierre Bourdieu: *Eine illegitime Kunst. Die sozialen Gebrauchsweisen der Fotografie*. Frankfurt / Main 1983. [GVK record]
- Manfred Brauneck: *Die Welt als Bühne. Geschichte des europäischen Theaters*. 6 volumes. Stuttgart 1993–2007. Volume 5 (2007). [GVK record]
- Stefanie Diekmann: Binnenschauplätze. Einige Anmerkungen zu Theater und Fotografie. In: Lars Blunck (ed.): *Die fotografische Wirklichkeit. Inszenierung – Fiktion – Narration*. Bielefeld 2010, pp. 105–116. PDF. DOI: [10.1515/transcript.9783839413692.105](https://doi.org/10.1515/transcript.9783839413692.105)
- Burcu Dogramaci: *Fotografie der Performance. Live Art im Zeitalter ihrer Reproduzierbarkeit*. Paderborn 2018. URN: urn:nbn:de:101:1-2018112117473202855517
- Johanna Drucker: Humanities Approaches to Graphical Display. In: *Digital Humanities Quarterly* 5 (2011), no. 1. HTML. [online]
- Johanna Drucker: *Graphesis: Visual Forms of Knowledge Production*. Cambridge, US-MA 2014. [GVK record]
- Miguel Escobar Varela: *Theater as Data: Computational Journeys into Theater Research*. Ann Arbor, US-MI 2021. PDF. DOI: [10.3998/mpub.11667458](https://doi.org/10.3998/mpub.11667458)
- Erika Fischer-Lichte: *Ästhetik des Performativen*. Frankfurt / Main 2004. [GVK record]
- Tancredi Gusman: Between Evidence and Representation: A New Methodological Approach to the History of Performance Art and Its Documentation. In: *Contemporary Theatre Review* 29 (2019), no. 4, pp. 439–461. PDF. DOI: [10.1080/10486801.2019.1657104](https://doi.org/10.1080/10486801.2019.1657104)
- Ian Hacking: *Representing and Intervening. Introductory Topics in the Philosophy of Natural Science*. Cambridge, US-MA 1983. [GVK record]
- X. Y. Han / Vardan Papyan / Ellen Prokop / David Leigh Donoho / C. Richard Johnson, Jr.: Artificial Intelligence and Discovering the Digitized Photoarchive. In: Lise Jaillant (ed.): *Archives, Access and Artificial Intelligence. Working with Born-Digital and Digitised Archival Collections (= Digital Humanities Research, 2)*. Bielefeld 2022, pp. 29–60. PDF. DOI: [10.14361/9783839455845-002](https://doi.org/10.14361/9783839455845-002)
- Donna Haraway: Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. In: *Feminist Studies* 14 (1988), no. 3, pp. 575–599. [GVK record]
- Sandra G. Harding: *The Science Question in Feminism*. 5<sup>th</sup> edition. Ithaca, US-NY 1993. [GVK record]
- Leonardo Impett / Franco Moretti: Totentanz. Operationalizing Aby Warburg's Pathosformeln. In: *New Left Review*, no. 107 (2017), pp. 68–97. DOI: [10.64590/nx9](https://doi.org/10.64590/nx9)
- Leonardo Impett / Fabian Offert: There Is a Digital Art History. In: *Visual Resources* 38 (2022), no. 2, pp. 186–209. HTML. DOI: [10.1080/01973762.2024.2362466](https://doi.org/10.1080/01973762.2024.2362466)
- Lise Jaillant (ed.): *Archives, Access and Artificial Intelligence: Working with Born-Digital and Digitised Archival Collections (= Digital Humanities Research, 2)*. Bielefeld 2022. PDF. [online]
- John D. Kelleher / Brendan Tierney: *Data Science (= The MIT Press Essential Knowledge Series)*. Cambridge, US-MA 2018. [GVK record]
- Friedrich Adolf Kittler: *Aufschreibesysteme 1800 / 1900*. 2<sup>nd</sup> edition. München 1987. [GVK record]
- Siegfried Kracauer: *Die Fotografie*. In: Siegfried Kracauer: *Das Ornament der Masse. Essays*. Frankfurt / Main 1977, pp. 21–39. [GVK record]
- Bruno Latour: *Reassembling the Social. An Introduction to Actor-Network-Theory (= Clarendon Lectures in Management Studies)*. Oxford#2005. [GVK record]
- Bruno Latour: *Pandora's Hope. Essays on the Reality of Science Studies*. Cambridge, US-MA 1999. [GVK record]
- Matthew Lincoln / Julia Corrin / Emily Davis / Scott B. Weingart: CAMPI: Computer-Aided Metadata Generation for Photo Archives Initiative. Whitepaper on Success and Failures in Computer Vision for Visual Digital Collections. KiltHub. 09.10.2020. PDF. DOI: [10.1184/R1/12791807](https://doi.org/10.1184/R1/12791807)
- Matthew Loper / Naureen Mahmood / Javier Romero / Gerard Pons-Moll / Michael J. Black: SMPL: A Skinned Multi-Person Linear Model. In: *ACM Transactions on Graphics* 34 (2015), no. 6. DOI: [10.1145/2816795.2818013](https://doi.org/10.1145/2816795.2818013)
- Lev Manovich: *Cultural Analytics*. Cambridge, US-MA 2020. [GVK record]
- Chantal Meyer-Plantureux: *La photographie de théâtre ou la mémoire de l'éphémère*. Paris 1992. [GVK record]
- Franco Moretti: *Distant Reading*. Konstanz 2016. [GVK record]
- Franco Moretti: ›Operationalizing‹: Or, the Function of Measurement in Literary Theory. In: *New Left Review*, no. 84 (2013). DOI: [10.64590/daw](https://doi.org/10.64590/daw)
- Franco Moretti (2000a): The Slaughterhouse of Literature. In: *Modern Language Quarterly* 61 (2000), no. 1, pp. 207–228. [GVK record]
- Franco Moretti (2000b): Conjectures on World Literature. In: *New Left Review*, no. 1 (2000). DOI: [10.64590/hxj](https://doi.org/10.64590/hxj)

- Ulf Otto: Post-Performance: Pandemic Breach Experiments, Big Theatre Data, and the Ends of Theory. In: Theater Research International (TRI) 48 (2023), no. 1, pp. 24–37. DOI: [10.1017/S0307883322000372](https://doi.org/10.1017/S0307883322000372)
- Ulf Otto: Plädoyer für eine symmetrische Theaterforschung. Über methodische Kälte, ethnografische Versuchungen und Lehren aus den Science und Technology Studies. In: Benjamin Wihstutz / Benjamin Hoesch (eds.): Neue Methoden der Theaterwissenschaft. Bielefeld 2020, pp. 247–270. DOI: [10.14361/9783839452905-012](https://doi.org/10.14361/9783839452905-012)
- Georgios Pavlakos / Nikos Kolotouros / Kostas Daniilidis: TexturePose: Supervising Human Mesh Estimation With Texture Consistency. In: 2019 International Conference on Computer Vision. 27 October – 2 November 2019. Proceedings. Los Alamitos, US-CA etc. 2019, pp. 803–812. DOI: [10.1109/ICCV.2019.00089](https://doi.org/10.1109/ICCV.2019.00089)
- Peggy Phelan: Unmarked. The Politics of Performance. London 1993. [GVK record]
- Matthew Reason: Documentation, Disappearance and the Representation of Live Performance. Basingstoke, UK 2006. [GVK record]
- John Resig: Using Computer Vision to Increase the Research Potential of Photo Archives. In: Journal of Digital Humanities 3 (2014), no. 2. HTML. [online]
- Hans-Jörg Rheinberger: Experimentalsysteme und epistemische Dinge. Eine Geschichte der Proteinsynthese im Reagenzglas. Göttingen 2006. [GVK record]
- Stefanie Schneider / Ricarda Vollmer: Poses of People in Art: A Dataset for Human Pose Estimation in Digital Art History. In: Journal on Computing and Cultural Heritage 17 (2024), no. 4. DOI: [10.1145/3696455](https://doi.org/10.1145/3696455)
- Benoît Laurent Auguste Seguin: Making Large Art Historical Photo Archives Searchable. PhD thesis, EPFL. 2018. DOI: [10.5075/epfl-thesis-8857](https://doi.org/10.5075/epfl-thesis-8857)
- Anastasis Stathopoulos / Ligong Han / Dimitris Metaxas: Score-Guided Diffusion for 3D Human Recovery. arXiv. 14.03.2024. DOI: [10.48550/arXiv.2403.09623](https://doi.org/10.48550/arXiv.2403.09623)
- Benjamin Wihstutz / Benjamin Hoesch (eds.): Neue Methoden der Theaterwissenschaft. Bielefeld 2020. DOI: [10.14361/9783839452905](https://doi.org/10.14361/9783839452905)
- Isa Wortelkamp: Bilder von Bewegung: Tanzfotografie der Moderne. Ilmtal-Weinstraße 2022. [GVK record]
- Phillip B. Zarrilli / Bruce McConachie / Gary Jay Williams / Carol Fisher Sorgenfrei (eds.): Theatre Histories: An Introduction. 2<sup>nd</sup> edition. New York etc. 2010. [GVK record]

## List of Figures

- Figure 1: Tryout of Detectron2 / OpenCV object detection (left), 2D pose estimation (middle), and image segmentation (right). [Production: *Die Weber*, Volksbühne Berlin 1993, director: Frank Castorf, photographs © Iko Freese]
- Figure 2: Example of scenic constellation in plain photograph (left), annotated by object detection (middle), and reduced by mesh recognition (right). [Algorithm ScoreHMR, production: *Anatomie Titus Fall of Rome*, Deutsches Theater 2007, director: Dimiter Gotscheff, photographs © Iko Freese]
- Figure 3: Bar charts of the shot times in the sample set of photos relating to *Anatomie Titus Fall of Rome* in photographer Iko Freese's private archive. This conveys not only an idea of the photographic practice (left), but also shows what moments in the play stick out and are repeatedly pictured (right). [EXIF metadata extracted and visualized with Python: Ulf Otto 2026]
- Figure 4: Visualization of all photos taken of the *Anatomie Titus* production in order of time taken (left), reduced to body meshes (right). [Human mesh extraction with help of ScoreHMR: Ulf Otto 2026]
- Figure 5: Misgendering is a common issue when dealing with theater data, which can be taken as a starting point for an exploration of gender construction in 20<sup>th</sup> century performing arts. [Screenshot of Detectron2 with OpenCV object detection on video recording of *The Persians*, Salzburger Festspiele 2018, director: Uli Rasche]
- Figure 6: Confidence intervals in face recognition can help to understand staging patterns: here, the thinking male in the front, accompanied by the caring female on the second layer. [Screenshot of Detectron2 object detection on video recording of *Ivanov*, Volksbühne Berlin 2014, director: Dimiter Gotscheff]