Exploring the *Exactitudes* Portrait Series with Restricted Boltzmann Machines

Sam D. Verkoelen, Maarten H. Lamers, Peter van der Putten

Media Technology group, Leiden Institute of Advanced Computer Science,
Leiden University, The Netherlands

sdverkoelen@gmail.com, {m.h.lamers,
p.w.h.van.der.putten}@liacs.leidenuniv.nl

Abstract. In this paper we explore the use of deep neural networks to analyze semi-structured series of artworks. We train stacked Restricted Boltzmann Machines on the Exactitudes collection of photo series, and use this to understand the relationship between works and series, uncover underlying features and dimensions, and generate new images. The projection of the series on the two major decorrelated features (PCA on top of Boltzmann features) results in a visualization that clearly reflects the semi structured nature of the photos series, although the original features provide better classification results when assigning photographs to series. This work provides a useful case example of understanding structure that is uncovered by deep neural networks, as well as a tool to analyze the underlying structure of a collection of visual artworks, as a very first step towards a robot curator.

1 Introduction

Exactitudes is an ongoing photography project by artists Ari Versluis and Ellie Uyttenbroek that started in 1993 and aims to document subcultures omnipresent in society ([14] of which a 6th ed. appeared in 2014). A sample of each subculture is captured in a series of twelve photographic portraits, all systematically taken from the same angle, with a neutral background and similar body position (Fig. 1). At present the project contains 154 series of twelve color photographs each, taken in the period 1994-2014.

The artists themselves appointed in an interview [9] that it "should almost be a scientific anthropological record of people's attempts to distinguish themselves from others by assuming a group identity". Although the portraits were intentionally presented by Versluis and Uyttenbroek within an arts context, they hold potential value for academics. In a study of sensory experience and affect in relation to denim clothing [1] the author compares different ways of wearing jeans, using Exactitudes as one of the resources. Other studies have discussed Exactitudes in the context of design [5], design teaching [12], fashion [6,11], and societal changes in consumption, activities and household organization [2].

Our work aims to explore the potential of applying deep learning neural networks to highly structured sets of visual artworks, and uses the Exactitudes collection of photographs as an example exploration. Primarily, we are interested in uncovering shared visual structure and variations thereon in sets of visual artworks, in relation to meta-features of these works. For example, do variations among a collection of visual artworks co-vary over time, or do sub-classes within a larger collection of visual

Figure 1. Four example *Exactitudes* series [14]. Left to right: 4 (*Bimbos*), 41 (*Surfistas*), 100 (*Cocktails & Dreams*), 154 (*United Americans*).

artworks correlate with combinations of visual features and meta-variables?

Clearly, the Exactitudes collection of photographs is visually highly structured. Within each series (twelve photographs depicting a subculture) the subjects wear clothing that identifies their sharing of a group identity, and the body position and stance are uniformly fixed. Moreover, visual identification of gender and (often) race and age are fixed within series. Across different series clothing, gender, race and body position/stance vary. However, overall position of the subject's body within the frame appears stable with few exceptions: most series depict subjects as full upper body, starting between hip and knee height, but five series deviate from this. Ultimately, all portraits are of equal horizontal and vertical dimensions and all backgrounds are uniformly grey/white, lacking visual features.

Our approach adheres to a common computational intelligence paradigm of viewing individual images, i.c. portraits, as separate points in a high-dimensional "pixel space". The many dimensions of this space each express a pixel position (from a digital representation) of the portraits, and span the possible values that each pixel can have. Through non-linear dimensionality reduction [3,8], we attempt to describe the individual portraits in much fewer dimensions, whilst as much as possible maintaining their underlying topographic relations in high-dimensional pixel space. The newly constructed dimensions in effect are non-linear combinations of the original pixel space dimensions, combined such that shared (non-linear) co-relations between the original dimensions in the data are encoded in fewer shared new dimensions.

A mapping of vectors in pixel space (i.e. images) to vectors in low-dimensional feature space, and v.v., are constructed within the resulting encoder. In effect, detailed information concerning shared visual structures in the portraits is captured inside the encoder tool, whereas individual portrait variations thereon are expressed as vectors spanning the lower-dimensional feature space. A deep neural network, in the form of stacked Restricted Boltzmann Machines [8], makes up our encoder.

In an exploratory manner and under the paradigm of discovery science [4], this study attempts to uncover as much information as possible about shared visual features among the Exactitudes portraits. It does so by scrutinizing the workings of the encoder, whilst treating it as a "black box" input-output processing device (c.f. [3]). In particular, we interpolate and extrapolate portraits by decoding positions in feature space to which none of the original Exactitudes portraits was precisely mapped. This method will expose both shared visual structures captured within the encoder, in particular their mutual non-linear dependencies and the individual variations thereon captured in portrait encodings. Subsequently, the individual portrait encodings will be viewed in relation to their meta-variables, such as series membership, so as to extract meaning from the encoding.

Through this process, we create knowledge on the application of deep learning neural networks to highly structured sets of visual artworks. Potential applications of this knowledge include in-depth discovery and understanding of structure in visual collections, automated curatorship of visual collections, automated authorship identification through outlier detection in lower-dimensional representations of visual works, and generative construction of new within-class visual works. Undoubtedly other potential applications exist.

In the remainder of this paper, we describe the methods applied, focusing more on the data set and the construction of our encoder tool. Subsequently, a subset from a larger series of small experiments [13] is described, along with their outcomes. [1] Finally, results and implications are reflected upon in a wider context of applying computational intelligence to visual arts.

2 Turning Exactitudes into Data

Digital portrait files were gathered with permission from the artists from their dedicated Exactitudes website (www.exactitudes.com). Structured filenames allowed for simple web-scraping techniques to collect all 1848 image files (154 series of 12 portraits), each a 600x600 pixel JPEG compressed RGB image with no alpha channel.

Five series of which the overall body framing of the portraits deviated significantly from the remaining 149 series were removed from further consideration (1, 7, 16, 18, and 77 using original numbering). Despite the uniform nature of the data, some further pre-processing was undertaken as to facilitate learning of the data by a deep learning neural network. Digital portraits were resized to 50x50 pixels and converted to 8-bit greyscale. Although this step is irreversible, it allows our experiments to be executed within a feasible timespan. Subsequently, pixel intensities were scaled to range [0,1] since our neural network model requires input values within this range.

Sparse representations of data have a number of benefits for energy based learning models as Restricted Boltzmann Machines [10], such as increased likelihood of correct classification. Sparsity i.c. means that the input vector representations consist mainly of zero values, with a limited number of non-zero values. A histogram of the greyscale pixel values shows that our data is not sparse. To magnify sparsity of the

[1] Generated experimental images, including and extending beyond those in this paper, are available for download at high resolution from https://samverkoelen.com/evomusart17/.

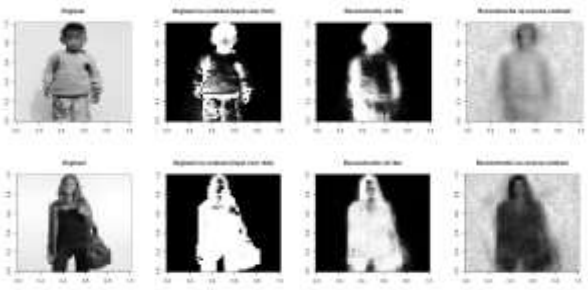

Figure 2. Two selected portraits. Left to right: (1) normalized, (2) contrast enhanced, (3) reconstructed from RBM^{50}, and (4) contrast enhancement reversed.

data, pixel values $x \in [0,1]$ were inverted after applying a sigmoidal contrast enhancing function in which gain g was set to 40 and cutoff c to 0.25, both values determined experimentally:

$$invcontrast(x) = 1 - \left(\frac{1}{1 + e^{g(c-x)}}\right)$$

This inverting contrast enhancement function creates sparse binary-like data vectors and can be reversed without any loss of information through applying $invcontrast^{-1}(x)$. What remains of our data is a sparse image space ($\mathbb{R}^{2500}$) containing 1788 representations of Exactitudes portraits.

3 Deep Learning Stacked Restricted Boltzmann Machines

We apply auto-encoding neural network models [3,8] for non-linearly reducing the dimensionality of our data. Specifically, our model consists of stacked Restricted Boltzmann Machines (RBMs), as proposed by Hinton [8]. Multiple RBMs are stacked on top of each other, where the feature layer (a.k.a. hidden layer) of a lower RBM provides inputs to the visible layer of the next-higher RBM. Moving upward, the number of feature layer units in each subsequent RBM is gradually reduced, and through an unsupervised contrastive divergence learning procedure this "deep" neural network stack is trained layer by layer.

Through pilot experimentation applying different numbers and sizes of RBM layers, it was assessed that five stacked RBMs were suitable for our data. In five steps, they transform the sparse image data from $\mathbb{R}^{2500}$ to feature vectors in $\mathbb{R}^{4000}$, $\mathbb{R}^{2000}$, $\mathbb{R}^{1000}$, $\mathbb{R}^{500}$ and finally $\mathbb{R}^{50}$. The initial "broadening" of the input vectors from $\mathbb{R}^{2500}$ to $\mathbb{R}^{4000}$ is required to capture the subtle non-binary character of the input data (with values approaching 0 and 1, but not limited to 0 and 1) in a larger subsequent binary layer [8]. The process of propagating data "upward" through the RBM stack transforming it from $\mathbb{R}^{2500}$ to $\mathbb{R}^{50}$ is referred to as "encoding". The resulting feature space is referred to as "RBM^{50}".

This architecture was selected based on visual assessment of the images' reconstruction quality for different architecture variations. Image reconstructions are created by propagating encoded RBM^{50} vectors "downward" through the RBM stack, expanding to $\mathbb{R}^{2500}$ and subsequently pixel-wise applying $invcontrast^{-1}(x)$, a process

Figure 3. Four linear connecting paths between two distant (D)RBM^{50} vectors.

henceforth referred to as "decoding". Since the mentioned reduction of Exactitudes portrait sizes and color information is irreversible, decoded (a.k.a. reconstructed) images are 50x50 pixels and in greyscale tone.

Fig. 2 illustrates a selection of input images and their reconstructions. A loss of visual information after RBM^{50} encoding/decoding of images is apparent and expected. Nonetheless, the human shape, posture, body position and luminosity of clothing are retained.

A principal component analysis was applied to the 1788 vectors encoded in RBM^{50}. Mapping RBM^{50} vectors onto their ordered eigenvectors is an optional final computational step of our encoder that is lossless and reversible. Although this step does not further reduce the dimensionality of data, it provides (linear) decorrelation of RBM^{50} vectors and ordering of features by variance in the data explained. The encoder including this final step is named the "decorrelated encoder". The resulting feature space is referred to as "DRBM^{50}". Dimensions of (D)RBM^{50} are referred to as "features".

4 Exploration through Experimentation

Under the paradigm of discovery science we subject the resulting encoder to a series of smaller experiments. These collectively form the exploration phase of our study, in which we attempt to uncover information about shared visual features among the Exactitudes portraits, and individual or grouped variation thereon.

Every experiment is described following the structure: method, results and observations. In this workflow, the results of one experiment may lead to a hypothesis that is tested in a following experiment. For sake of brevity, we can only provide selected images resulting from experiments. Larger versions and collections of experimental images are available from [13].

4.1 Straight paths between feature vectors pairs

Method. For the 50 most distant pairs of (D)RBM^{50} portrait vectors, a reconstruction of their straight connecting path in (D)RBM^{50} is made in 10 equidistant steps. It is

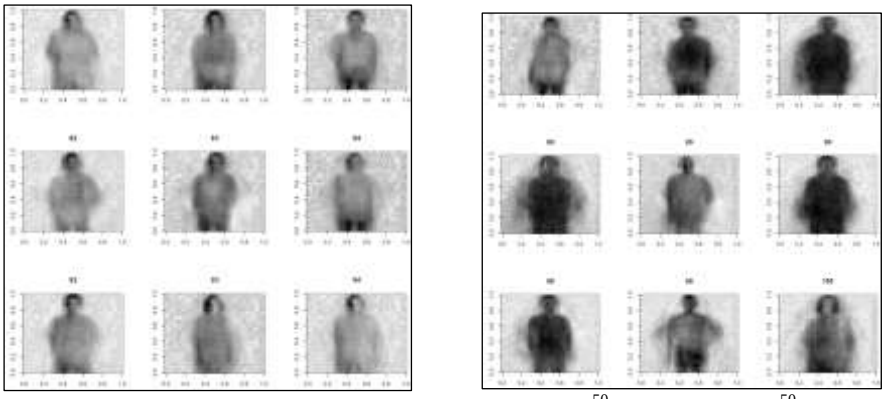

Figure 4. Reconstruction from nine random points in RBM^{50} (left) and $DRBM^{50}$ (right)

noteworthy to mention that the most distant pairs in RBM^{50} do not necessarily have great distance in $DRBM^{50}$.

Results. From the resulting 50 paths, a subset is chosen based on visual disparity, and shown in Fig. 3 (top two in RBM^{50}, bottom two in $DRBM^{50}$).

Observations. The outermost reconstructions from RBM^{50} (top two) show a clear visual difference in color intensity, posture and body position. The outer reconstructions from $DRBM^{50}$ (bottom two) have less visual differences, despite their relative large Euclidean distance. This indicates that distance in RBM^{50} is more related to visual differences, as opposed to distance in $DRBM^{50}$. Regardless of the position in either RBM^{50} or $DRBM^{50}$, each intermediate vector on a connecting line reconstructs to an image that could be assessed as human-shaped.

4.2 Random points in feature space

Method. From the previous experiment one might hypothesize that every arbitrary point in (D)RBM^{50} reconstructs into an image with human shape. Here, 100 random points (c.f. a uniform probability distribution) in (D)RBM^{50} are reconstructed into images for visual review.

Results. A random selection is presented in Fig. 4.

Observations. The hypothesis that every arbitrary point in (D)RBM^{50} reconstructs into and image with recognizable human shape is visually confirmed. It is notable that the reconstructions' color intensity differs between RBM^{50} and $DRBM^{50}$ vectors, the latter appearing noticeably darker.

The reconstructions indicate that decorrelation changed the distribution of portrait vectors across RBM^{50}, placing darker photos throughout $DRBM^{50}$ whilst the lighter ones are closer together. This observation might also explain the lack of visual differences in $DRBM^{50}$ reconstructions in Fig. 3.

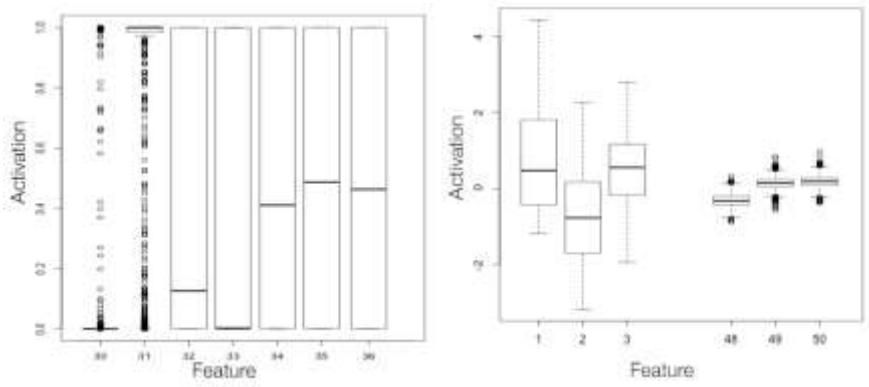

Figure 5. Box plots of RBM^{50} (left) and DRBM^{50} (right) activations over the entire dataset.

4.3 Activation distributions per feature

Method. Activations per (D)RBM^{50} feature are summarized in box plot visualizations, over the entire dataset (n = 1788).

Results. Features of DRBM^{50} are ordered by the variance each one explains of the portrait collection. Thus the first explains most, subsequently decreasing until the least explanatory feature. Fig. 5 (right) shows activations of the first and last three features of DRBM^{50}. In RBM^{50} there is no particular ordering of features; a random subset of 7 features and their activation distributions are shown in Fig. 5 (left).

Activations are considered an outlier if the value is outside 1.5 times the interquartile range above the upper quartile or below the lower quartile. Outlier positions are marked with a small circle in both box plots.

Observations. As expected, lower-numbered DRBM^{50} features show more variance than higher-numbered ones. Portrait activations in RBM^{50} are distributed differently because the feature order is arbitrary, and many differences exists (Fig. 5, left). For example feature 30 is fairly outspoken at zero with a few exceptions, whereas feature 34 is about similar in high and low activations.

Activations in DRBM^{50} are mostly around the center whereas activations in RBM^{50} are mostly located on the outer edges. For the latter, a potential explanation is the probabilistic nature of Restricted Boltzmann Machines [8] in combination with contrast enhancement.

4.4 Highest and lowest activating portraits per feature

Method. As observed in experiment 4.3, feature activations in RBM^{50} are disordered and can be outspoken, contrasting DRBM^{50} in which features are ordered and activations congregate around center-values. We now order all portraits by their activation on every individual feature in (D)RBM^{50}, revealing the highest and lowest activating portraits per feature.

Figure 6. Highest and lowest activating portraits on RBM^{50} features 30 and 34.

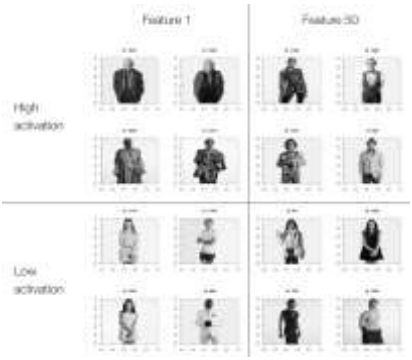

Figure 7. Highest and lowest activating portraits on DRBM^{50} features 1 and 50.

Results. Fig. 6 illustrates the highest and lowest activating portraits for RBM^{50} features 30 and 34. These features were selected based on results of experiment 4.3. Fig. 7 illustrates the same for DRBM^{50} features 1 and 50.

Observations. Fig. 6 shows that outspoken activations observed in experiment 4.3 on RBM^{50} feature 30 correspond to apparent visual differences in portraits, i.c. mainly the color intensity. In contrast, less outspoken RBM^{50} feature 34 correlates with less salient visual differences. Fig. 7 shows the highest and lowest activating portraits on the first DRBM^{50} feature (left), exposing a correlation with both color intensity and posture. Less outspoken DRBM^{50} feature 50 (right) correlates to less visually salient differences.

4.5 Single feature variations in portrait context

Method. To visualize effects of feature changes on portrait reconstructions, every (D)RBM^{50} feature is individually set to its lowest and highest observed value over the dataset. All other feature activations are set according to an existing portrait, providing context to the visual assessment of reconstructions.

Results. The complete results consist of the high and low feature variations, for every (D)RBM^{50} feature on a randomly selected subset of 10 portraits [13]. Fig. 8 shows variations for features 30-35 on a single portrait.

Observations. RBM^{50} reconstructions in Fig. 8 lack apparent visual diversity, whereas in DRBM^{50} they show visual diversity not only across features 30-35, but also across low and high activations. E.g. DRBM^{50} feature 30 correlates to color intensity, and 34 to apparent arms position.

4.6 Single activation feature vectors

Method. Encouraged by the results of experiment 4.5, here all (D)RBM^{50} features are set to their minimum possible value (zero) with the exception of a single highlighted

feature being set to its maximum value (one). Reconstructions were made as such for all features once highlighted (50 in RBM^{50}, 50 in $DRBM^{50}$). To visually enhance differences between reconstructions, their mean reconstruction (n = 50) was subtracted from each individual reconstruction.

Results. A single visually salient feature was selected and presented in Fig. 9. High scoring areas in the reconstructions appear as lighter color intensities, and v.v.

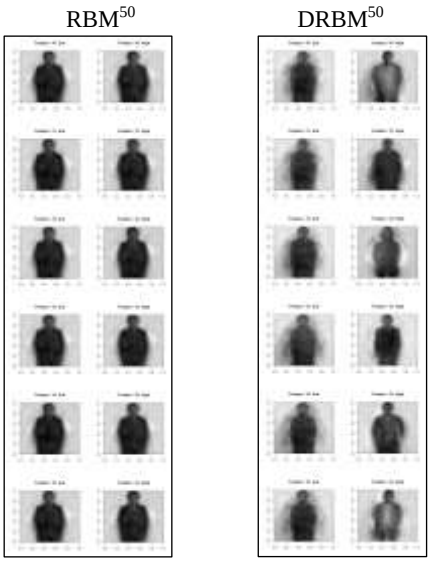

Figure 8. Variations on one portrait, setting features 30-35 (top to bottom) to their lowest (left) and highest (right) activations, in RBM^{50} and $DRBM^{50}$.

.

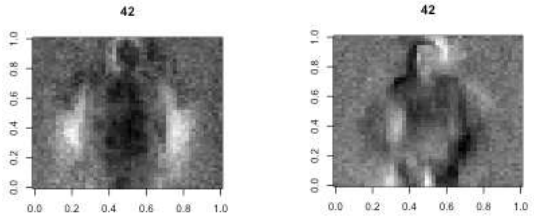

Figure 9. Reconstructions of feature vectors containing all zeros and value 1 on feature 42, in RBM^{50} (left) and $DRBM^{50}$ (right).

Figure 10. Highest and lowest activating portraits on RBM^{50} and $DRBM^{50}$ features 42.

Observations. Reconstructions on all highlighted features show high and low scoring areas on various body parts [13]. Visually apparent in the RBM^{50} reconstruction of Fig. 9 are the high scoring areas around the arm positions, whereas high scoring regions in the $DRBM^{50}$ reconstruction are more "spread out". This topological salience is also apparent from portraits that activate low and high on RBM^{50} feature 42 (Fig. 10), although it is less so in $DRBM^{50}$ where there is arguably and in general a lesser visual relationship between high scoring reconstruction areas and high and low activating portraits.

4.7 Portrait distributions over feature space

Method. Visualizing $(D)RBM^{50}$ is hardly possible in a meaningful manner. However, the distribution of portrait encodings over $(D)RBM^{50}$ can be essential information, e.g. for training a classifier. Although it captures not all information, a histogram of the Euclidian distance from the center of $(D)RBM^{50}$ for all 1788 portrait encodings, provides a rough impression.

Results. Fig. 11 shows the described histograms.

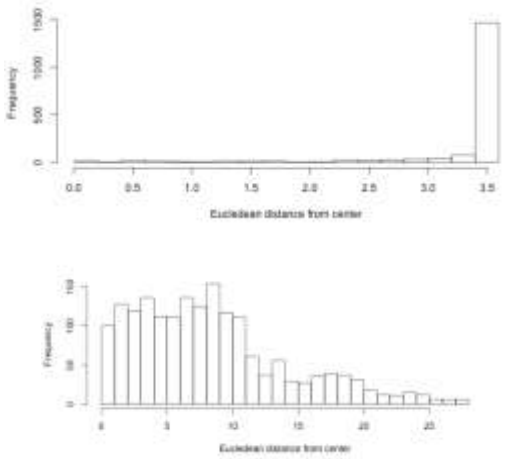

Figure 11. Distribution of portrait encodings' Euclidian distance to the center of RBM^{50} (top) and $DRBM^{50}$ (bottom); n = 1788.

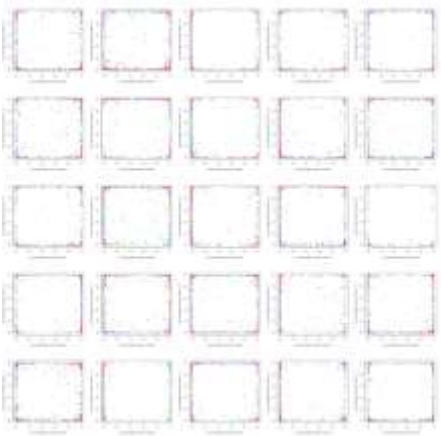

Figure 12. Portrait activations on feature pairs in RBM^{50}.

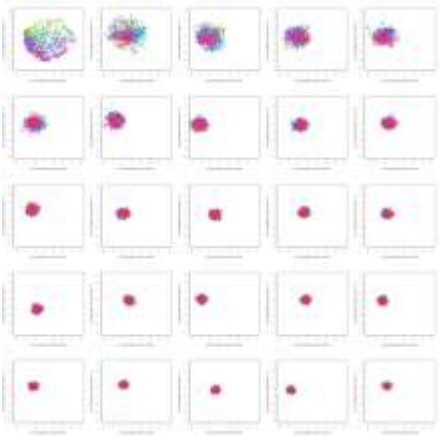

Figure 13. Portrait activations on feature pairs in DRBM^{50}.

Observations. Distributions of portrait encodings across RBM^{50} and DRBM^{50} are quite different, as to be expected. Observations from experiment 4.2 and 4.3 are confirmed by these results: portrait encodings in RBM^{50} are mostly situated on the outer edges, whereas they are situated more around the center of DRBM^{50}.

4.8 Feature pair visualizations

Method. In an attempt to further understand the structure of (D)RBM^{50}, portrait activation values (n = 1788) on pairs of features are plotted, while color-coding for series (n = 149).

Results. Plots of portrait activations on feature pairs {1,2}, {3,4}, {5,6}, ... , {49,50} in RBM^{50} (Fig. 12) and DRBM^{50} (Fig. 13) feature space are made. Plots are ordered row by row, from left to right, and generated with the lowest feature number on their horizontal axis. Larger versions in [13]. Series membership is coded with random color assignment.

Observations. Portrait encodings in RBM^{50} appear to have no particular ordering, but show a conspicuous pattern where activation values are distinctively along the space's edges (Fig. 12). Portrait encodings in DRBM^{50} form a circular clustered shape (Fig. 13). These observations are c.f. experiments 4.2, 4.3, and 4.7. Plots in Fig. 13 are clearly ordered by variance explained per feature, an effect that is also apparent from the slight horizontally oriented oval point clouds in plots.

Series are marked individually with a different color each, something that is not visually apparent among 149 series (Figs. 12 and 13). A smaller subset of series plotted onto DRBM^{50} features pair {1,2} are visually much more distinct and clearly clustered (Fig. 14). When augmenting the same plot with originating portraits (Fig. 15), the arrangement of portraits appears structured along visual characteristics. DRBM^{50} features pair {1,2} is able to not only separate series, but also visually distinct portraits within series, e.g. spacing light and dark toned 'Topshoppers' apart.

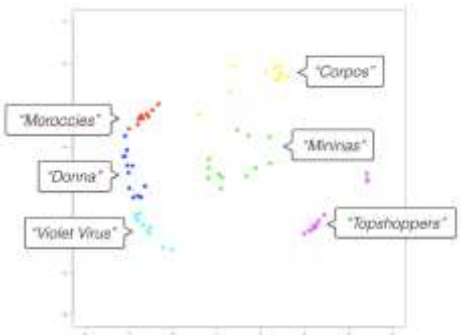

Figure 14. Portrait activations on DRBM^{50} features pair {1,2} for six labeled series.

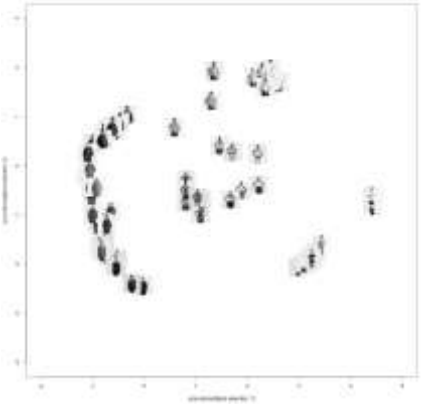

Figure 15. Idem Fig. 14, labeled with originating portraits.

4.9 Series classification

Method. From Figs. 14 and 15, questions arise regarding series classification from portrait encodings. A random forest classifier [8] is repeatedly trained and tested on randomly resampled training (1500) and test data (288). Results for three encodings are compared: RBM^{50}, DRBM^{50}, and mapping onto 50 principal components of the normalized portraits (referred to as PC^{50}).

Results. Classification correctness rates are reported in Table 1.

Figure 16. Five best (left) and worst (right) classified series, each displayed horizontally.

	RBM^{50}	**DRBM^{50}**	**PC^{50}**
Trial 1	0.622	0.517	0.642
Trial 2	0.615	0.531	0.677
Trial 3	0.628	0.535	0.667
Mean	0.622	0.528	0.662

Table 1. Classification correctness rates after different encodings over three trials.

Observations. All three methods are clearly successful in that they far outperform hypothetical 149-sided die classification (expected correctness 0.007). Nevertheless it appears that decorrelation of RBM^{50} features has negative effect on the ability to classify encoded portraits by series. Intuitively, this contrasts the visual distinctiveness that decorrelation appears to have added, c.f. Figs. 12-15. This observation is even more striking when noting that PCA by itself (in the form of PC^{50} encoding) outperforms RBM^{50} in terms of classification, albeit only marginally.

4.10 Best and worst classifiable portraits

Method. The relatively poor contribution of DRBM^{50} to series classification appears to discord with prior expectations. Attempting to gain insight into the classification process, the best and worst classified series are inspected.

Results. Notably, the five best and worst classified series were the same under both RBM^{50} and DRBM^{50} encoding, and illustrated in Figs. 16 and 17 resp.

Observations. The corresponding results for RBM^{50} and DRBM^{50} indicate that the classifier likely learns the same differences between series. This observation leaves uncertain why the classification results between both feature encodings differ. Notable also is the homogeneity within the best classified series. Although homogeneity is a characteristic of Exactitudes, it is to a greater or lesser degree present within individual series. As expected, the least classifiable series appear more heterogeneous.

5 Discussion and Future Work

In this concluding section we reflect on the experimental results, and introduce the robot curator as future potential application area for this type of research.

5.1 Analytical reflections

Despite their popularity and interest from researchers, dimensionality reduction methods such as deep learning are commonly referred to as black boxes. This exploratory study aimed at giving an insight in the inner workings of such networks when applied to semi-structured portrait photos of the Exactitudes project. Understanding is gained using an exploratory approach, one might consider atypical: given a dimensionality reducing network, referred to as *encoder*, a series of small experiments was conducted. These experiments were of manageable size, making it possible to develop new experiments based on previously obtained results and explore encoders from different viewpoints.

This approach provided a number of noteworthy results that will be reflected upon. However, the encoder itself has a notable aspect also, namely the use of a sigmoid-shaped contrast enhancement function to force the input data into sparse binary-like vectors. Prior research did not indicate a similar approach to handle real valued data in Restricted Boltzmann Machines. Our results showed that the network could learn the binary-like data, whilst being able to reverse the transformation and reconstruct input images. Although our study only uses photos as network input, there is no reason to suspect that this approach is limited to this class of data.

Novel results emerged from the experiments. For example experiments 4.3 and 4.4 together suggest a relation between the outspokenness of a feature and the visual differences it is able to explain, wherein outspoken means that the average feature activation on the entire dataset is near one or zero. The more outspoken a feature is, the more visual discrepancy it can explain.

Furthermore, the optional decorrelation of feature vectors showed varying results. Experiments 4.3, 4.7, and 4.8 clearly visualize a transformation of feature space, as to be expected. However, due to this transformation, features that were distilled in RBM^{50} feature space are intertwined in the decorrelated $DRBM^{50}$ feature space, as experiment 4.6 indicated. In this experiment a division is made on both posture and intensity, making it more difficult to select photos on an individual feature.

Decorrelation also showed an advantage, by visually enhancing the separation between series (Fig. 14), whilst simultaneously discriminating on visual differences (Fig. 15). These results make it worthwhile to use feature decorrelation when eyeballing visualizations for clusters or groups within the dataset. However, a distinction should be made between the human eye and classification algorithms. Experiment 4.9 shows that feature decorrelation adversely effects series classification, something that is difficult to explain from a computational viewpoint.

The same experiment also indicates that series classification on the first 50 principal components of the normalized portraits outperforms the previous methods. Although all classifications outperform random classification by a roughly 85 times

higher success rate, it is noteworthy that a relatively simple technique such as principal component analysis can compete with contemporary dimension reduction algorithms. Despite the enormous popularity of dimension reduction or deep learning, it remains worthwhile to make a comparison with well-established techniques.

Furthermore, what this research has yielded in a broader sense is the approach itself. Because of its exploratory nature, the path to follow is less clear in comparison with more traditional hypothesis-driven studies. Thus, there is an increased need for structure and guidance. Performing small and manageable experiments that lead to new hypotheses and experiments in rapid succession is, to our best knowledge, novel for exploratory research into deep learning.

The method allows new insights to be fitted directly into new experiments, thus adjusting the direction of the research based on preliminary results and advancing insights. One must ensure in such exploratory research that the problem is viewed from multiple angles, something we have done to our very best in this study.

When looking back at the exploratory goal, we find this research has made a positive contribution to the understanding of deep neural networks. By means of various visualizations, a better understanding of what is learned in these networks was obtained. This understanding could not only help in the selection of relevant features but also bring light to possible shortcomings in the data set or previously unknown features that might be beneficial. Therefore a good understanding of what is learned adds considerable value in exploring and exploiting large datasets.

5.2 The robot curator: artificial curation as a computational creativity task

In our introduction we identified prior research in arts and humanities as well as a range of potential future applications of our research. In this section we propose one specific area for future application: automated curation.

When creating an exhibition, choices are made about what works to include, and also in what context to present these works, for example by spatial organization. Hence curation is not just a value free reflection on works of the artists but can also been seen as a form of artistic expression in itself. The curator aims to tell a story not limited by chronological order only, may want to make the public aware of hidden relationships between works within an oeuvre or across artists, provide new ways of seeing connections and relationships or actually want to juxtapose works that each provide a different perspective to create dialogues or interventions.

Given that curation can be seen as a creative activity, it is a potential task of interest within computational creativity research. This can be approached as an artificial creativity task; both fully automated curation as well as curation augmentation, where man and machine have a more symbiotic role — comparable with artificial intelligence versus intelligence augmentation.

Our study can be seen as a simplified example of how computational learning techniques can be applied towards this aim. The various features provide new ways to organize Exactitudes series and individual portraits, or to provide a spatial ordering of series that could be leveraged for a real world exposition. For example series could be organized on a floor plan according to their topology in decorrelated feature space. Likewise, the encoder could be used to augment existing artworks with new portraits, through various generative methods given.

This is mostly barren territory for research and many directions could be explored. How can we create methods that are suitable within a semi-structured body of work by a single artist, but also generalize across artists? Moreover, there is no reason to restrict this approach to visual art only, or to artists working within a single domain. One could focus on just photography or paintings, but could also take multimodal information such as background texts, artist network structures or other background knowledge into account. Naturally, for augmented creativity approaches, the robot curator will need to make space for letting the human curator into the loop.

Acknowledgements

We acknowledge Ari Versluis and Ellie Uyttenbroek as the creators of Exactitudes and owners of all intellectual rights and privileges, and are grateful to them for sharing this wonderful collection with the world via www.exactitudes.com.

6 References

1. Candy, F.J. (2005). The fabric of society: an investigation of the emotional and sensory experience of wearing denim clothing. *Sociological Research Online*, 10(1)
2. Casimir, G. (2011). Interaction of Societal Development and Communication Technology. *International Journal of Home Economics* 4(1), pp. 3-13
3. DeMers, D., Cottrell, G. (1993). Non-linear dimensionality reduction. *Advances in Neural Information Processing Systems* 5, pp. 580-587
4. Doherty, P.C. (2008). The Beginner's Guide to Winning the Nobel Prize: Advice for Young Scientists. Columbia University Press
5. Gardien, P., Djajadiningrat, T., Hummels, C., Brombacher, A. (2014). Changing your Hammer: The Implications of Paradigmatic Innovation for Design Practice. *International Journal of Design*, 8(2), pp. 119-139
6. Grinde Arntzen, M. (2015). *Dress Code: The Naked Truth about Fashion*. Reaktion Books
7. Han, J., Kamber, M., Pei, J. (2011). *Data mining: concepts and techniques*. Elsevier
8. Hinton, G.E., Salakhutdinov, R.R. (2006). Reducing the dimensionality of data with neural networks. *Science* 313, pp. 504-507
9. Huffington Post. (2012) *Artists Create Anthropological Photo Series*. Available from: www.huffingtonpost.com/2012/06/25/exactitudes-interview-art_n_1619483.html. Retrieved on: 2016-11-10
10. Ranzato, M.A., Poultney, C., Chopra, S., LeCun, Y. (2006). Efficient learning of sparse representations with an energy-based model. *Advances in Neural Information Processing Systems*, 19, pp. 1137-1144
11. Smelik, A. (2011). The Performance of Authenticity. *ADDRESS Journal for Fashion Writing and Criticism,* 1(1), pp. 76-82
12. Trocchianesi, R., Guglielmetti, I. (2012). Design teaching and cultural companies: Languages, tools and methods toward a profiable involvement. *Strategic Design Research Journal*, 5(1), pp. 49-57
13. Sam D. Verkoelen, (2015). Exploring Dimensionality Reduction on Semi-structured Photos — A Closer Look at Exactitudes. *Master's Thesis for the Media Technology programme, Leiden University (The Netherlands)*
14. Versluis, A., Uyttenbroek, E. (2002). *Exactitudes*, 010 Publishers, Rotterdam