Art as Data Set, Data Set as Art: Training machine vision systems to see as artists

Daniel Buzzo
Creative Technology Lab,
University of the West of England
Coldharbour Lane, Bristol, UK
daniel.buzzo@uwe.ac.uk

The process of an artist 'speaking' to an audience by illustrating meaning underlying a body of work on a subject is commonplace within contemporary art. In the age of mechanical reproduction artists can quickly and simply assemble and collate coherent sets of visual materials as a means of non-verbal investigation. In the world of technical images (Flusser 2011) the meaning in a photograph or technically produced image is buried beneath the surface. Artists often assemble bodies of work, as collections of related materials or multiples, sometimes in the style of the visual essay. This allows the artist to illustrate, through type and repetition and focus and observation, that which is important in the pattern of things to an audience. This process assists the audience to see 'what relates to what' between items in the body of work in a non-verbal visual, semantic or cognitive way. This collation and relation of pattern, or similarity, of defining an unspoken meaning with exemplar work is also the foundation of how computer vision systems classify and 'understand' the world. This classification comes from human 'trainers' assembling materials as data sets and ascribing meaning to the data sets in an ontological way. In this paper I describe the process of creating a data set for machine learning from a photographic collection. The photographic materials collected as part of a larger artwork. In the text I explain how this dataset is an extension of the artwork and also an artwork in it's own right. As datasets used for training machine vision systems are commonly centrally generated for commercial/industrial processes I argue that the idea of an artist generated artwork /dataset is the beginning of a process of taking control of the means of surveillance, teaching machine vision systems to see not just how the surveillance capital commercial hegemony sees but how artists see. This creation of artist's data sets and data sets as art is in itself an intervention as art work. We are proposing an evolution of the arguments originating in the industrial revolution of seizing control of the means of production. Now in the age of the surveillance society and surveillance capitalism artists are proposing 'Seizing control of the means of surveillance'.

Surveillance capitalism. Machine learning. Imaging Photographic essay. Data activism.

1. INTRODUCTION

The pursuit of visual art has often been a mission of the capturing, recording and re-presentation of non-verbal intention by artists to the wider public. A process of portraying to an audience ideas, concepts and observations that cannot, or possibly will not, be satisfactorily described through other means of communication. This portrayal and 'cognitive communication' of an underlying or non-explicit meaning takes place in the world post what Flusser would call 'Technical Images'. These are part of a visual system that has often relied on bodies of works, collections, repetitions and series to portray and describe the specific meanings or observations underlying the surfaces of an image. (Flusser 1983) This mode of presenting numerous examples of what the artist sees is a process of illustrating their point of attention she hopes to educate or train an audience to notice or understand. That which is important or of note in the images is conveying or emphasising that which is common between the presented collections of works. Toister asks in the paper 'Why be a Photograph?" what is the unique epistemic advantage of photography, and confronts the anxiety about the 'end' of photography. (Toister 2014)

> 'The photographic images we are working with today, and the ones that will be available to us in the near future... will only be understood within a theoretical context that describes photography as ***an extreme form of mathematical abstraction***.' (emphasis added)

Figure 1: CCTV warning sign on a shelf front amongst dental hygiene products, Bristol, UK.

Toister also echoes Flusser when thinking of the photograph as explicitly a carrier of encoded information. Where the photographer is facilitating the camera to make models of the world. In that the camera-as-system is making ontological decisions, categorisations, and meaning-structures of both what-can-be-seen and also what-can-be-known of the world.

> 'What do I actually mean when I say a photograph of a house depicts that house, and a computer image of an airplane yet to be built is a model? … Any way I formulate the difference between depiction and model, I come to grief … It can therefore be said of a photographer that he has made a model of a house in the same sense that the computer operator has made a model of a virtual airplane.' (Flusser 2011, pp.42–43).

When taking the creation of sets of materials that illustrate a type or particular meaning to a human audience one can draw direct parallels with the current process of training machine vision systems to recognise objects and patterns. In this process the creator of the dataset (artist) collates visual materials that illustrate the underlying concept or object that he or she wishes the receiver (e.g. classifier or neural network) to 'comprehend', 'understand' or recognise. The receiving system (deep learning or human audience) is entrusted with the task of discerning a common pattern or underlying shared significance within the dataset (visual essay or body of work). After exposure to the training data set (exhibited body of work) it can successfully repeat or recognise new materials external to the original dataset (body of artwork) that exhibit similar traits or characteristics. In this

sense it can be argued that, in the same way that an artist seeks to alter the minds or perceptions of an audience the dataset seeks to alter the underlying structure of a recognition system or network. Sometimes this desire to reconfigure or retrain perception, awareness and categorization in an audience by an artist comes in radical forms. In the words of American conceptual artists Lawrence Weiner (Wiener 2017, Ainley 2017)

> 'That's the whole point of what already is, already is an objectification of something that you are trying to communicate with a culture. It's that tiny little thing that you don't want to fuck up someone's day but the purpose of what you are doing is to fuck up their whole life.'

2. SIGNS OF SURVEILLANCE: THE PROJECT

The underlying project that is the source materials for this artist data-set began as an initial observation and then a non-verbal investigation of the phenomena of growing public surveillance. Across Europe and the wider globe, surveillance of the body public, of civic space, of every interaction in everyday society appears at some kind of saturation point.

Figure 2: CCTV warning sign from Stockholm, Sweden. ©The Author.

In its 2013 report, 'The Picture Is Not Clear: How Many CCTV Surveillance Cameras Are There in the UK?' the British Security Industry Association (BSIA) suggested that camera numbers in the private sector could outnumber those used by public bodies by as much as 70 to 1. The BSIA survey covered the whole of the UK, not just London, and its maximum estimate suggested there was a CCTV camera for every 11 people in the country.

In 2015 the then UK Information Commissioner, Tony Porter, himself a former Counter-Terrorism officer said...

Figure 3. 'Signs of Surveillance' Interactive Web and animated video Installation artwork ©The Author.

'The lack of public awareness about the nature of surveillance troubles me'

...and as Matthew Thomas wrote in the Guardian newspaper, quoting Porter; (Weaver 2015):

'[it] risks changing the "psyche of the community" by reducing individuals to trackable numbers in a database'.

In 2004 Richard Thomas the then UK Government Information Commissioner warned (Booth 2004):

'My anxiety is that we don't sleepwalk into a surveillance society where much more information is collected about people, accessible to far more people shared across many more boundaries, than British society would feel comfortable with'

Making a comparison to the collection of vast quantities of information on individual of Franco's fascist state and the eastern European communist regimes, he added in response to the question of whether there was a risk of Britain following this route:

'I think there is a danger. I don't think people have woken up to what lies behind this. It enables the government... to build up quite a comprehensive picture about many of your activities.'

Against this backdrop, and as part of several activities and outputs as part of the is project the dataset 'Signs of Surveillance' was created explicitly for the purpose of training computer vision systems. Systems trained using this data-set can subsequently recognise places where computer vision systems or other kinds of video surveillance may be occurring. The contents of the dataset are based upon a six year project photographing, collecting and collating these signs. The collection is taken from public spaces across the world, indicating that CCTV surveillance is operating. (see Figure 1, 2) These systems, and their associated warning signs, increasingly indicate that automated, trained machine vision systems such as ANPR (automated number plate recognition) and face recognition, amongst other systems, are being used. The original photographic collection contains over 2000 images recorded in 15 different countries and includes images from Belgium, Canada, China, Denmark, France, Italy, Germany, Japan, Luxemburg, Netherlands, Portugal, Sweden, UAE, United Kingdom and United States. In addition to releasing the source images formatted as an image training data set for machine learning the project has produced an interactive web map and animated video installation showing the source and location of the images on a navigable, searchable world map (see Figure 3). This portion of the project as more conventional interactive artwork is documented in detail in (Buzzo 2020).

2.1 Process and method and material

The corpus of images contains photographic material illustrating signs of surveillance – principally signs warning the public of the use of CCTV – from fifteen countries, – each image has been separated by country and processed and formatted in preparation for use in training deep learning machine vision systems. The data set is an adjunct and addition to existing datasets for training such as the extensive ImageNet (http://www.image-net.org/) database.

From a data-science rather than from an art perspective the image set is principally designed for studying the problem of unconstrained CCTV sign recognition in the wild. Each image has been prepared by hand to create a uniform set of consistent aspect ratio and pixel data per image. Ready for passing to the selected neural network model for training. In common with other image sets such as 'Labelled faces in the Wild' (http://vis-www.cs.umass.edu/lfw/#download) et al the images are available to download as a processed 250x250 pixel square set ready for training and also as the un-cropped high resolution set for alternative uses and cross-testing. [9] With this image-data set artists and researchers using computer vision, machine learning and image based neural network technologies can train their respective systems to be able to automatically detect the existence of CCTV 'Signs of Surveillance' in the wild.

2.2 Making 'critical' datasets as activist resistance

The specific point of this work, from an artistic and activist perspective is to illustrate clearly how the nature of data used to train future 'sentiences' or in this case automated recognition and classification systems is entirely dependent upon the data supplied to train them. At present all data supplied is commonly from central sources of what Dunne and Raby in the now (in)famous 'Critical Design FAQ' might have called 'affirmative' datasets, that is data sets that confirm and re-inforce the status quo of hegemonic and societal classification of the world. (Dunne and Raby 2007) These affirmative datasets are as opposed to 'critical' data sets, ones that challenge the existing elite or predominant ontological vision of the world.
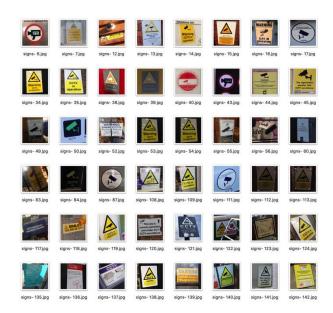

Figure 4: *Section of images from the UK portion of training images in first version of 'Signs of Surveillance' dataset. ©The Author*

This first 'Critical Data set' is aimed squarely at the military/industrial/surveillance capital model that creates data sets of things that are ontologically useful or are desirable or assumed in a hegemonistic sense. As with ideas of Critical Design, it is proposed that data sets can be also be divided into two distinct types, the first is 'Affirmative', that is, dataset that re-inforce the worldview, values and status quo of the financial and political interests that fund and pursue the development of automated surveillance and image recognition systems. The alternative data set is 'Critical' in that it explicitly questions the nature, ontology, world-view and purpose of existing 'Affirmative' data sets.

As early artistic forays into this area explicitly acknowledged such as the work of the Surveillance Camera Players and guides to the *Guerilla Programming of Video Surveillance Equipment*. (Surveillance Camera Players 2006) (Carter 1995) where Carter says:

> 'It is important to remind oneself of the relationship between the eye of the media and that of the corporate police state – for they are both the guardian of the commodity, however nebulous and ephemeral that commodity may become. As a tactic designed to point out the paradox of a system that turns the lens on a public that has been taught to place more importance on images recorded by cameras than images seen by their own eyes, we propose *Guerilla Programming of Video Surveillance Equipment* (emphasis added).'

To this end, the 'Signs of Surveillance' data set can be explicitly used to create automated surveillance systems that can detect the presence of other automated surveillance systems. In this sense it is a 'critical' data set. This act of artists and non-technocrats creating datasets with which to train the machines and systems that have become some of the foundations of what is referred to as the system of 'surveillance capitalism' is to be seen in the context of radical thinkers or writers such as Newton or DuBois. (Newton 1969, DuBois 2013) where they argued for the seizing of the means of production we now are in the midst of a post-industrial struggle to **'Seize the Means of Surveillance'.**

3. CONCLUSION

The dataset (a small sub-section of which is illustrated in Figure 4) is being shared as an open source image training set under a creative commons (attribution) license for non-commercial use in machine vision training projects. This dataset comprises a 48Mb.zip format archive and contains 1,830 prepared and categorised items. It is presented as both body of artwork and as dataset for activism and is the first in a series of similar works aimed explicitly at taking an active hand in shaping globally deployed systems.

The dataset and instructions for its use is now available freely to download from the 'Signs of Surveillance' project page at http:/signsofsurveillance.com. The images and dataset are retained as copyright of the author but are made available and free for non-commercial use with share-alike attribution.

4. REFERENCES

Booth, J. (2004) 'UK "sleepwalking into Stasi state"', *The Guardian*, 16 August. Available at: https://www.theguardian.com/uk/2004/aug/16/britishidentity.freedomofinformation.

British Security Industry Association: (2013) The Picture is Not Clear: How many CCTV surveillance cameras in the UK. Tech. Rep. 195, British Security Industry Association, Worcester, UK. URL https://www.bsia.co.uk/publications/publications-search-results/195-   the-picture-is-not-clear-how-many-cctv-surveillance-cameras-in-the-uk.aspx (retrieved 25 April 2011).

Buzzo, D. (2020) 'Signs of Surveillance' in *"The Arts, Design and Technology"* Editors: Rae Earnshaw, Susan Liggett. Springer.

Carter, M. (1995) *Guerilla Programming of Video Surveillance Equipment*. Available at: http://www.notbored.org/gpvse.html (Accessed: 1 July 2019).

Dunne, A. and Raby, F. (2007) *Critical Design FAQ*, *Dunne and Raby*. Available at: http://www.dunneandraby.co.uk/content/bydandr/13/0 (Accessed: 21 March 2017).

Flusser, (1983). V. *'Towards a philosophy of photography'*, *English*, pp. 1–96.

Flusser, V. (2011) *Into the Universe of Technical Images*. Edited by K. Hayles, M. Poster, and S. Weber. University of Minnesota Press.

Huang, G. B. et al. (2007) *Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments.*

Huey P. Newton, (1969) *Ebony,* (August 1969).

Lawrence Weiner. *(2017) Art Should Fuck Up Your Life: The Zen of Lawrence Weiner*. VICE magazine. Available at: https://video.vice.com/en_us/video/art-should-fuck-up-your-life-the-zen-of-lawrence-weiner/58f53c132574666019e2c21b

Nathaniel Ainley (2017) *Lawrence Weiner Told Us About F**king Up Your Life With Art*, *Vice*. Available at: https://www.vice.com/en_us/article/wjqgkz/lawrence-weiner-fcking-up-your-life-with-art (Accessed: 1 December 2019).

The Surveillance Camera Players (2006) *Surveillance Camera Players:10-Year Report*. Available at: http://www.notbored.org/10-year-report.html (Accessed: 6 July 2019).

Toister, Y. (2014) 'Why be a photographic image ?', *Philosophy of Photography*, 5(2), pp. 161–167.

Weaver, M. (2015) 'UK public must wake up to risks of CCTV, says surveillance commissioner', *The Guardian*. Available at: http://www.theguardian.com/world/2015/jan/06/tony-porter-surveillance-commissioner-risk-cctv-public-transparent.

W.E.B DuBois. (2013) *The Autobiography of W. E. B. DuBois*. Diasporic Africa Press.