Bayesian Networks for fMRI Analysis

Nicole Ostlund (s1755552), Ina Quataert (s1135546),

Meeke Rijnen (s1856723), Marieke Vinkenoog (s1136305)

Introduction

Image recognition and classification in humans is almost instantaneous: show someone a picture of a cat and they'll immediately recognize it as one. However, ask them to explain why this is a cat, and not for instance a dog, and they probably won't have an answer ready. Both have four legs and a tail, and are furry, but even without consciously knowing the exact criteria we use to distinguish them, we have no trouble discerning cats from dogs. This classification problem has led researchers to a swath of models and tools to teach computers to learn the parameters that would easily reconcile differences for a human.

In this essay, we will focus on using Bayesian networks for fMRI data. First, we will explain how Bayesian networks can be used to classify images. Next, we will give a short introduction to fMRI. After that we will combine the two previous sections and combine Bayesian network image classification with fMRI pictures. Moreover, we will show that Bayesian networks not only can be used to classify images but also for decoding images. Lastly, we will end with a conclusion where we will give our opinion on the subject.

Using Bayesian networks to classify images

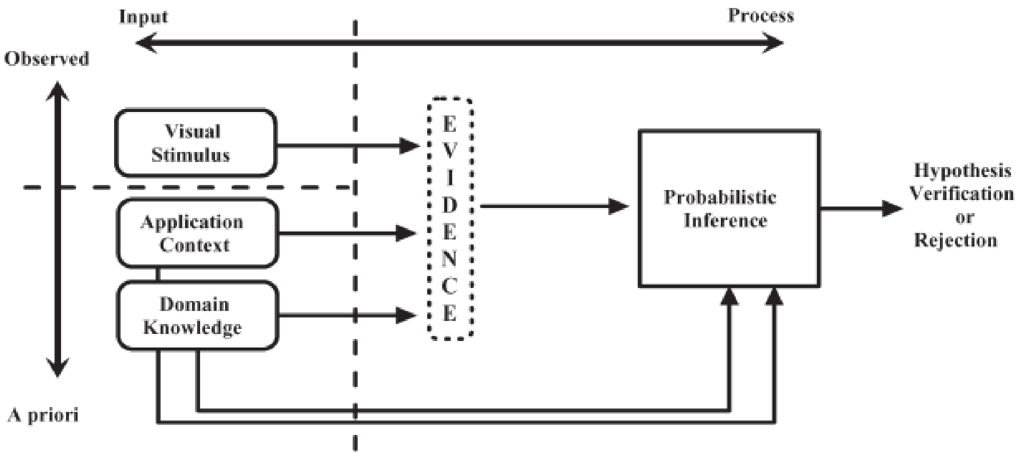
There are several ways to approach computer vision techniques, the two most popular types being neural networks and Bayesian networks. Neural networks are often a sort of black box; it is not completely clear how weights and nodes are determined, while Bayesian networks give more insight in the features of the image used to classify it. Both can be learned by a computer from data and achieve very high accuracies. Bayesian networks also combine expert or prior knowledge with the evidence observed from the image which allows for a greater utilization of domain knowledge and puts given data in a specific context. Bayesian networks used for image recognition make use of the famous Bayes Rule:

$$P(S|I) = \frac{P(S) * P(I|S)}{P(I)}$$

Where p(I|S) is the image likelihood (viewpoint, lightning), P(S) is the prior regularities (geometry and shape, material, lightning and expert knowledge) and P(I) is the image regularities (geometric and photometric properties).

This framework, proposed by Nikolopoulos et al. (as seen in Figure 1), makes the procedure of classification and interpretation more explicit than other networking algorithms.

Figure 1. Bayesian inference in image classification as proposed by Nikolopoulos et al.



To use this framework in image categorization, a set of formulated hypotheses is tested. Each category concept, c_p refers to a hypothesis that best describes and image, I_q. Therefore, the hypotheses are $H(I_q) = \{\Pr(I_q) : i = 1, \dots \dots, n\}$, where *n* is the number of category concepts.

First, a global classifier is applied to the image to help determine the initial probability for each hypothesis. The global classifier, based on a model that is trained using global image information, gives some context for the image. The type of image determines what parts of the image should be looked at in more detail by a local classifier. For instance, if it is known from the global classifier that we're looking at a picture of a forest landscape, there's not as much use for trying to find a car in the image as there would be if we were looking at a photograph of a city.

After the global classifier determines which evidence needs to be taken from the image, local classifiers are applied to segments of the image. The local classifiers generate a set of confidence intervals that we then take as the evidence:

$$E(I_q) = \{\Pr(\acute{c}_i | I_q^{\delta j}) : i = i, \dots k, j = 1 \dots . m\}$$

where $I_q^{\delta j}$ are the image regions, $\acute{c}_i$ the regional concepts, *k* the number of regional concepts and *m* the number of identified segments. Next, localized region labelling is applied. This is the task of assigning labels to each image region, choosing from one of the regional concepts. Again, we formulate a hypothesis for each image segment:

$$H(I_q) = \{\Pr(I_q^{\delta j}) : i = 1 \dots k, j = 1 \dots . m\}$$

where k and m are the same as before. Conflicts can occur in this framework, when the local classifier suggests a concept that does not belong to the set of concepts returned by the global classifier. In that case, a choice must be made which classifier to believe. It is debatable which classifier is more reliable, and this should be decided by the researcher.

Figure 2. Application of classifying procedure to an image, demonstrating the passing of information between local classifiers

	Global Classifiers
Tennis	45,97
Roadside	54,21
Rockyside	47,07
Seaside	**56,31**
Forest	52,46
C.Buildings	56,00

	Local Classifiers			
	Region1	Region2	Region3	Region4
Board	48,51	49,51	50,62	46,84
Court	49,72	52,53	51,83	54,00
Person	52,56	50,16	51,34	52,93
Gradin	49,07	53,01	51,40	51,67
Court line	50,13	49,25	51,85	51,93
Racket	48,93	50,99	50,26	48,34
Sky	47,03	**60,69**	**74,54**	**56,03**
Dried plant	47,14	49,63	48,09	47,97
Grass	**56,36**	53,29	47,66	50,31
Road	52,35	48,47	49,45	54,97
Vegetation	49,17	47,18	46,85	53,74
Ground	50,55	48,56	50,70	49,51
Road line	48,55	49,93	49,89	50,10
Car	47,83	49,51	47,60	47,92
Trunk	48,82	47,68	49,00	48,41
Rock	49,13	47,40	47,79	48,00
Glacier	50,00	50,00	50,00	50,00
Sea	48,73	51,32	47,87	44,08
Sand	49,62	47,86	49,66	47,41
Wave	52,85	46,54	47,96	48,29
Boat	49,76	49,74	47,63	48,77
Tree	50,47	47,61	48,41	48,25
Stone	49,58	49,50	49,39	49,55
Building	44,62	47,71	46,14	43,97
Roof	48,60	52,30	49,47	49,90

	Belief Evolution					
%	Tennis	Roadside	Rockyside	Seaside	Forest	C,Buildings
Prior Probabilities	15,4	17,6	15,7	17,3	**18,5**	15,4
Global	13,4	20,2	14,2	**21,2**	20,1	18,8
Evidence-Region1	13,4	**21,7**	14,2	21,2	20,9	20,6
Evidence-Region2	13,4	**23,5**	15,2	22,9	21,5	22,6
Evidence-Region3	13,4	**27,2**	17,3	26,3	22,6	26,8
Evidence-Region4	13,4	**27,2**	17,6	26,9	22,8	27,4

Figure 2 shows an example of how this framework is applied to a photograph. The belief evolution starts with the highest (prior) probability of a forest classification, because this is the category most present in the database used. After running a global classifier, the beliefs are updated. Now the highest scoring category is seaside, which is still incorrect. Then, local classifiers are applied to the four regions depicted in the image. The local classifiers correctly classify region 1 as grass and region 3 as sky, but misclassify regions 2 and 4 as sky also. Still, this information from the local classifiers is enough to shift the belief from seaside to roadside, which is the correct classification for this image.

Using a Bayesian network is particularly useful for image recognition since we can express the image in uncertainty and incorporate expert prior knowledge. The application to neuroimaging can be articulated deductively, as a mechanism of image processing to find pathways (using regional classifiers), and inductively, by using those pathways in scans to reconstruct what may have been seen. These Bayesian networks for neuroimaging make it possible for researchers to deduce mechanisms of relationships that were otherwise not well known, provide information to the operations of the complex human brain, and potentially improve the understanding of many psychiatric and neurological diseases such as Alzheimer's and Parkinson's diseases.

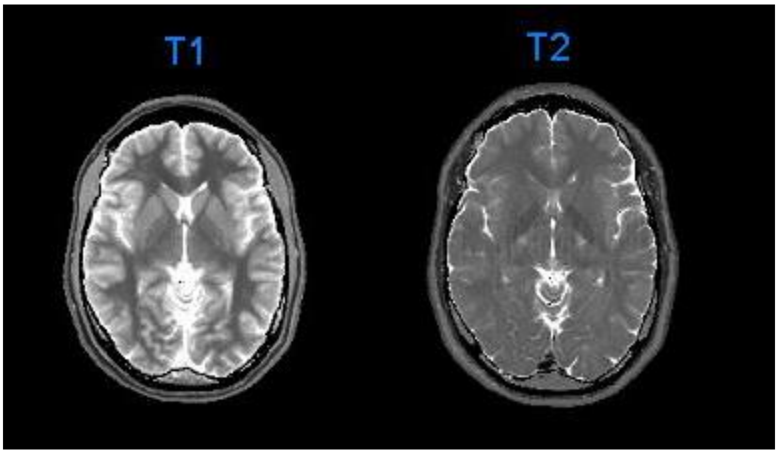
Introduction to fMRI

When applying image recognition to a processing of images related to brain and brain behavior, the procedural mechanisms to create the image should be well understood. In short, the fMRI uses a combination of magnetism and radio pulse technology to alter protons in the human body and capture the distortions in the magnetic field as a series of images. When a patient is placed inside the MRI machine, the protons in the hydrogen elements of water molecules in the body align with the direction of the magnetism. Radio impulses programmed at frequencies matching the spin of those protons are pulsed at different intervals. These pulses cause the protons to shift in some way (changes in spin, changes in direction, tilt, etc.) and align in phase. These protons are unstable in this phase and will quickly start to dephase back to the original alignment.

The changes in intensity of these phases are picked up as signals categorized by their time intervals: T1 (longitudinal relaxation) and T2 (Transverse Relaxation). After the RF pulse is removed, the magnetism of the protons that have been tilted decrease and this relaxation of the protons longitudinally (in the z direction) is caught in the T1 weighted images. The T2 signal on the other hand captures the change in magnetism along the transverse plane, where the protons dephase but do not move longitudinally (they stay on the X-Y plane). The T1 and T2 images are captured by manipulating the data collection time after the radiofrequency pulse.

The collections of images can be categorized in two ways by the two different resolution types: temporal resolution or spatial resolution (see Figure 3). The structural T1 image has high spatial resolution, but low temporal resolution; in other words it is a static image. This is the type of image used when brain structures or abnormal growths need to be distinguished. A low resolution T1 image is usually the type of image used to identify the anatomical characteristics in a subject before running the actual functional tests. This helps to orient both the technicians and the researchers. The functional T2* (which is a combination of the T2 and local inhomogeneities in the magnetic field) image makes use of temporal resolution at the expense of spatial resolution; the images are followed over time and space but the objects are not as easily distinguished in the image. The collection of T2* images of a voxel over time is called a time series of that voxel.

Figure 3. The two time-interval dependent captures, demonstrating the difference in characteristics of images.



The most common mechanism of fMRI voxel mapping is the Blood Oxygen Level-Dependent (BOLD) response. The magnetism of blood changes with oxygenation and deoxygenation of hemoglobin. Protons close to the deoxyhemoglobin spin at different speeds than those far away, which causes a different T2* gradient across voxels of high “activity.” The connection between deoxygenated blood and neural activity is not perfected, but, for posterity, active brain areas are defined as places in the brain with high concentrations of deoxygenated blood as it is a mark of that part of the brain tissue working. In effect, the higher metabolic demands of neurons, the higher deoxyhemoglobin in that area of the brain.

These images, specifically the T2 images as they inform both temporal and spatial understanding, and their serialization provide the groundwork for brain activity research.

Bayesian network and Image Classification of fMRI

As described, fMRI provides an indirect measure of brain activity by means of a blood-oxygenation-level-dependent (BOLD) signal. This characteristic makes it useful to apply a Bayesian network to analyze functional brain connectivity and the connectivity during cognitive challenges. Both constraint based and score based algorithms have been used for this data including PC, Casual PC, Greedy Equivalent Search, Cyclic Causal Discovery, and Fast Causal Inference (Mumford & Ramsey, 2014). The methods PC, CPC, CC, and FCI are constraint-based methods (make inferences from conditional independences), while GES is a score based method (model selection based on a scoring mechanism). In addition to these techniques, other techniques known as Dynamic Bayesian Networks, such as LiNGAM, LOFS, and GIMME, allow for temporal information to be incorporated and are better suited for detecting both connections and their directionality (Mumford & Ramsey, p. 574). For our purposes, we'll focus on the network types more broadly, with a specific emphasis on the properties of the Dynamic Bayesian Network, which has had demonstrated successes with fMRI analyses.

Modeling Techniques

When analyzing fMRI it is very common to select regions of interest (ROI's), which occur as a cluster of voxels, in order to study the functional integration (connectivity) between these ROI's. The ROI approach makes it possible to model the functional connectivity as a Bayesian network, where each ROI constitutes as a node and the functional connection between the ROI are represented by an arc of the graph. Moreover, the conditional distribution can be used to compute the strength of the connection. In Figure 4 the steps involved from original fMRI data to Bayesian network modelling can be seen. Important to note are two ways to enter the BOLD signal into a Bayesian network model:

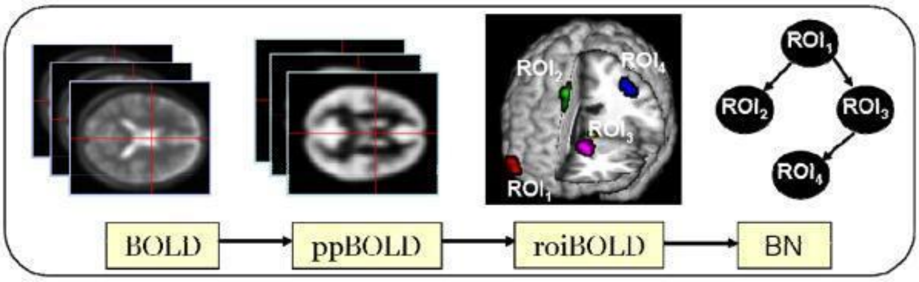
1) Discrete nodes: BOLD time series are discretized or quantized into a finite number of categories (nonlinear)
2) Linear Gaussian interactions: assumes that the neural interaction between regions can be modeled by a linear relationship with Gaussian noise.

In conjunction with these are two overall ways to model a Bayesian network that are relevant for fMRI data:

1) Static Bayesian networks: the functional network is considered static across time
2) Dynamic Bayesian network: functional network is considered dynamic and describes spatial and temporal relationships (time-series data). This is a more

realistic approach since the brain is a dynamic procedure and the images used capture temporal and spatial information.

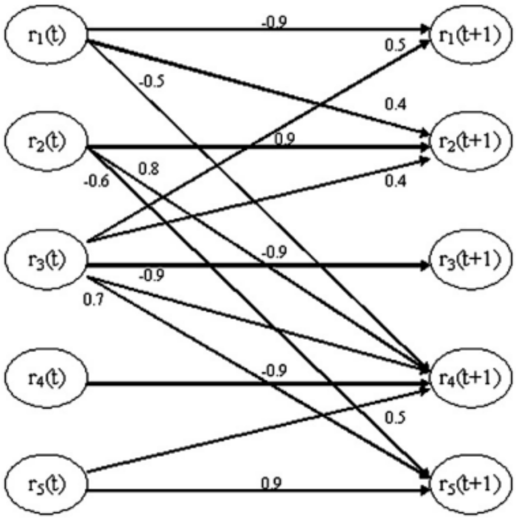
Figure 4. Summary of steps involved in connectivity analysis, from original fMRI data to Bayesian network modelling



An example of a dynamic Bayesian network of 5 ROI regions can be found in Figure 4. It is interesting to see what the differences are between static Bayesian networks and dynamic Bayesian networks. In an experiment, both networks found the same network structure. However, dynamic Bayesian networks are capable of learning the structure more accurately as they explicitly take into account the temporal characteristics of fMRI time-series by using a Markov chain (Rajapakse J.C & Zhuo J, 2007). Using a static Bayesian network results in a loss of information about edge directions because several network structures with the same skeleton but different edge directions can have the same marginal likelihood while a dynamic Bayesian network avoids this by taking the temporal relationships into account.

To conclude, for the most realistic results a dynamic Bayesian network should be used. There is only one drawback of the dynamic Bayesian network, which is that a large sample is needed. However, we still have a strong preference for the dynamic Bayesian network approach since it is most similar to how the real brain works. A dynamic Bayesian network procedure uses a Markov Chain with nonlinear continuous time interactions (instead treating the time intervals as discrete) and provides a direct mechanism to model temporal relationships between brain regions. The use of nonlinear interactions allows the assumption of linearity, which often does not hold, to be skirted making this method attractive when other methods have failed (Burge et al, 2009). An example network of 5 ROI regions can be found in Figure 5.

Figure 5. A dynamic Bayesian network representing a neural system consisting of five brain regions. The values of the edges represent the strength of connections.



Improving models

In order to improve accuracy of the models, images can be preprocessed. This procedure, known as highpass filtering, allows the noise of the images to be greatly reduced. By using the methods detailed below, which can either be used in addition to the original images to improve learning or replace the images entirely in cases of too much noise, the learning of image characteristics can be greatly improved.

1) *Alignment*: corrects for misalignment of the images across the slices and scan sessions, originating basically from the head movement. Assuming the head as a solid, rigid-body transformations (rotation and translation) are applied, based on some image similarity measures.

2) *Slice-timing correction*: when associating each slice (in-plane image) with a time point, it must be noted that there is a time delay from the start to the end of scanning of each single slice. This phase delay must be corrected, using some interpolation technique.

3) *Unwarping*: corrects for distortions in the images, generated by *inhomogeneities* in the magnetic field. In addition to the machine's imprecision, head size and location are sufficient to create inhomogeneity or *bias field*.

4) *Spatial normalization*: registers different subject's brains to a common *stereotaxic space*, such as *Talairach* space and Montreal Neurological Institute (MNI) space. Normalization is necessary when comparing subject's BOLD signals in a group.

Normalization also allows the localization of a particular brain structure of interest through an anatomical atlas.

5) *Smoothing*: a spatial filtering is performed by means of a Gaussian kernel to reduce noise and enhance the statistical power for group comparisons. This also has to do with the imprecise nature of the spatial normalization process.

Decoding-reconstruction using Bayesian networks.

In addition to the application of Bayesian networks to classify images, Bayesian networks can be used to in the reconstruction of images. Yargholi and Hossein-Zadeh investigated whether Bayesian networks could improve decoding-reconstruction. Decoding-reconstruction is used to reconstruct the actual visual stimuli (images) from measured brain activity patterns. The applied procedure for decoding-reconstruction is shown in Figure 6. For this study, fMRI data from Van Gerven et al. (2010) is used. The data consists of 100 trials from one subject. In each trial a hand-written digit of a 6 or a 9 on a black background was visually presented to the subject for 12.5 seconds. Brain activity patterns were measured and used to apply decoding-reconstruction.

A Decoding-Reconstruction Bayesian network

There are several steps to set up efficient Bayesian networks for decoding-reconstruction, which are now explained. First, nodes need to be determined using train data. Nodes are represented by brain voxels and pixels of reconstructed stimuli. To avoid overfitting of the data, the number of parameters and dimensions needed to be reduced. The number of parameters is reduced by converting the stimuli images into binary images and by putting a threshold on the activity of brain voxels so that they would have 2 states; active or inactive. The number of dimensions is reduced by sorting brain voxels and selecting the ten voxels that distinguish most between the two stimuli using the Kolmogoro-Smirnov test (KS-test). Also, the number of pixels are reduced using two methods: by using perpendicular blocks of 7 pixels and by using the ten pixels that distinguish best between the two stimuli using KS-test.

Next, there is the procedure of structure learning of the Bayesian networks. Effective connections between nodes were extracted using the search-and-score structure learning method Greedy Search. This resulted in two types of effective connections; pixel – voxel edges and edges between voxels. In total, six different Bayesian networks were constructed to use for decoding-reconstruction, existing from different partitioning of stimulus pixels and different types of edges (table 1). Additionally, a Greedy Search was used to learn connections between pixel – pixel edges. These connections were added to the basic Bayesian network structure that only consisted of pixel – voxel edges, formerly.

Table 1. The six learned Bayesian networks with different partitioning of stimulus pixels and different edges in the structures.

	Partitioning of stimulus pixels	Existing edges in BNs' structure		
		Pixel–voxel	Voxel–voxel	Pixel–pixel
BNs1	Perpendicular blocks of 7 pixels	√	–	–
BNs2		√	√	–
BNs3		√	–	√
BNs4	10 pixels in the *p*-value queue of the KS-test	√	–	–
BNs5		√	√	–
BNs6		√	–	√

Lastly, parameters for the Bayesian networks are learned using maximum likelihood estimation. Hereby, the influence of applying brain effective connectivity information on decoding-reconstruction was explored.

Influence of Bayesian networks on decoding-reconstruction

Each reconstruction was evaluated objectively and subjectively. Objective evaluation was done by deriving the city-block distance between stimulus (s) and its reconstruction (r), each having N pixels:

$$D(s,r) = \frac{1}{N}\sum_{i=1}^{N} |s_i - r_i|$$

Subjective evaluation included 3 persons grading the similarities between stimuli and their different reconstructions with a 1 (lowly), 2 (on average) or 3 (highly). Also, the performance of each Bayesian network is obtained by adding the scores and scaling them. This is called the winning percentage.

Table 2 shows separate comparisons between Bayesian networks 1–3 and between Bayesian networks 4–6. It is clear that Bayesian networks 2 and 5 are the winners, as they have the lowest city-block distance errors, the highest winning percentages and the highest subjective assessment scores compared to the other Bayesian networks.

Table 2. The average (standard deviation) of city-block distance errors, winning percentages and subjective assessment scores for Bayesian networks 1–6.

	BNs1	BNs2	BNs3
City-block distance error average(std)	0.1209(0.0079)	0.1035(0.0080)	0.1223(0.0084)
Winning percentage (%)	20	70	10
Subjective assessment scores (out of 100)	76	90	60

	BNs4	BNs5	BNs6
City-block distance error average(std)	0.1095(0.0080)	0.0944(0.0078)	0.1175(0.0079)
Winning percentage (%)	14	84	2
subjective assessment scores (out of 100)	83	100	76

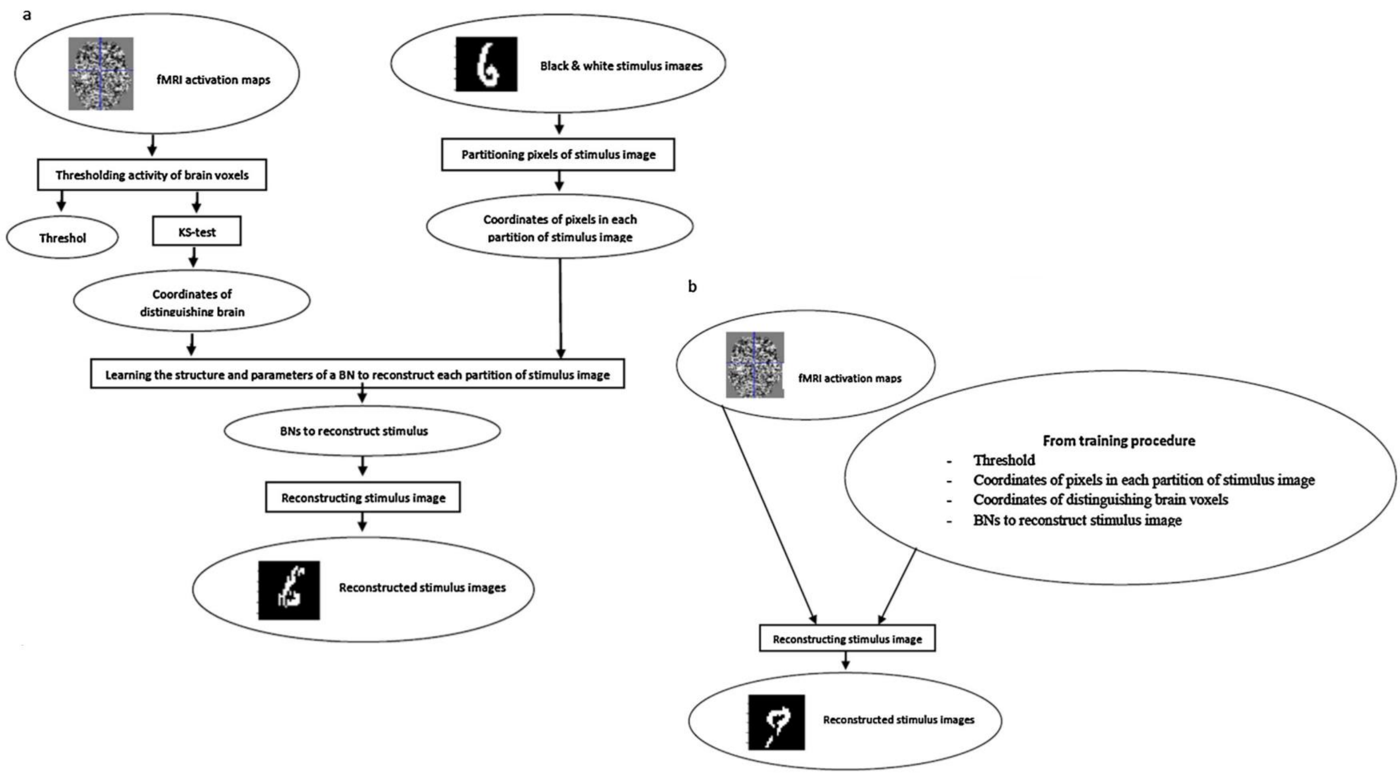
From a comparison between all networks, Bayesian network 5 turns out to be the overall winner (Table 3), as the winning percentage for this network is much higher compared to the other networks. Thus, a Bayesian network that uses the ten pixels that distinguish best between the two stimuli using KS-test and has pixel – voxel edges and edges between voxels performs best.

Table 3. The winning percentages for Bayesian networks 1–6.

	BNs1	BNs2	BNs3	BNs4	BNs5	BNs6
Winning percentage (%)	0	16	0	7	75	2

The results of this study reveal that Bayesian networks offer improvements to the decoding-reconstruction of handwritten digits. The Bayesian networks in this experiment demonstrate how different brain voxels, pixels from presented stimuli, and connections between them add information to the process of reconstructing a visual stimulus from brain activity patterns.

Figure 6. The (a) training procedure and (b) testing procedure for decoding-reconstruction.



Conclusion

In this paper, we investigated the use of Bayesian networks for fMRI data and showed some examples of how it is implemented. In our opinion, the main benefit of using Bayesian networks is knowing what is happening behind the scenes. Other methods, such as neural networks, which we've encountered in other courses, worked well in image classification tasks, but were not theoretically well-founded making thorough understanding of them difficult. With Bayesian network this problem is solved making the image classification easily retraceable and defensible. This makes it possible to easily see where the system was wrong or where it can be improved. However, we also believe that there is a drawback in this as it makes computations harder and it less automated.

Limitations

Though the use of Bayesian networks conceptually fits the research questions more appropriately than other methods, a meta-analysis found that for readily confronted problems in fMRI (directionality, session lengths, etc.) the Bayesian network approach failed when too many nodes were present (the computation became too expensive) and the use of LiNGAM, a non-linear approach, failed if too few nodes were present (Smith et al, 2011). As fMRI sessions are extremely variable, insensitivity to data size makes the use of Bayesian networks a bit problematic. Certain methods such as High Order Dynamic networks have already demonstrated an improvement and this problem will likely dissipate with increases in computational power (Yargholi & Hossein-Zadeh, 2016).

Additionally, when using Bayesian networks for fMRI processing the researchers must be careful in choosing the appropriate model, as the wrong model might not be able to capture directionality (typically Gaussian procedures) and depending on the level of inference. Certain algorithms, can only make conclusions on a group level (GIMME), while other algorithms can also be applied to single subjects, such as LiNGAM and LOFS (Mumford & Ramsey, 2014). This puts the burden on the researcher to be mindful of the types of conclusions that they would like to make.

In addition to these few procedural problems, conceptual decisions regarding priors and expert knowledge becomes increasingly important and the lack of or fallibility of priors could lead your Bayesian networks astray. This is probably why we have seen "Discussion" sections of papers repeatedly advising their successors to proceed with caution when making decisions for their own networks. Overall, however, as researchers often make careful choices in regards to research, the use of Bayesian networks for classifying images, specifically fMRI data, has proven to be very useful since it does not use only static calculations but can be made dynamic. Especially for the fMRI case this is convenient as fMRI is not static at all. Using Bayesian networks for decoding-reconstruction of fMRI data is relatively new, but proven to be efficient. Bayesian networks help by adding more structure in decoding-reconstruction.

References

Burge, J., Lane, T., Link, H., Qiu, S. and Clark, V. P. (2009), Discrete dynamic Bayesian network analysis of fMRI data. Hum.

Ide, J. S., Zhang, S., & Li, C. R. (2014). Bayesian network models in brain functional connectivity analysis. *International Journal of Approximate Reasoning : Official Publication of the North American Fuzzy Information Processing Society*, *56*(1 Pt 1), 10.1016/j.ijar.2013.03.013. http://doi.org/10.1016/j.ijar.2013.03.013

Mumford, J. A., & Ramsey, J. D. (2014). Bayesian networks for fMRI: A primer. *NeuroImage,86*, 573-582. doi:10.1016/j.neuroimage.2013.10.020

Nikolopoulos, S., Papadopoulos, G. T., Kompatsiaris, I., & Patras, I. (2011). Evidence-driven image interpretation by combining implicit and explicit knowledge in a Bayesian network. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 41*(5), 1366-1381.

Rajapakse J.C, Zhuo J (2007). Learning effective brain connectivity with dynamic Bayesian Network. *NeuroImage 37(749-760)*, http://cs.ru.nl/~peterl/BN/RajapakseZhou2007.pdf

Smith, S. M., Miller, K. L., Salimi-Khorshidi, G., Webster, M., Beckmann, C. F., Nichols, T. E.,Woolrich, M. W. (2011). Network modelling methods for FMRI. *NeuroImage,54*(2), 875-891. doi:10.1016/j.neuroimage.2010.08.063

van Gerven, M.A.J., de Lange, F.P., & Heskes, T. (2010). Neural Decoding with Hierarchical Generative Models. *Neural computation, 22*(12), 3127-42. doi: 10.1162/NECO_a_00047

Yargholi, E., & Hossein-Zadeh, G.A. (2016). Reconstruction of digit images from human brain fMRI activity through connectivity informed Bayesian networks. *Journal of Neuroscience Methods, 257*, 159-167. doi: https://doi.org/10.1016/j.jneumeth.2015.09.032