Title

Top-Down Modulation of Neural Responses in Visual Perception: A Computational Exploration

Permalink

https://escholarship.org/uc/item/8918v3p8

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 29(29)

ISSN

1069-7977

Authors

Mozer, Micheal C.
Fan, Adrian

Publication Date

2007

Peer reviewed

Top-Down Modulation of Neural Responses in Visual Perception:
A Computational Exploration

Michael C. Mozer and **Adrian Fan**

{michael.mozer,adrian.fan}@colorado.edu
Department of Computer Science and Institute of Cognitive Science
University of Colorado, Boulder, CO 80309-0430

Abstract

Visual perception is typically performed in the context of a task or goal. Nonetheless, visual processing has traditionally been conceptualized in terms of a fixed, task-independent hierarchy of feature detectors. We explore the computational implications of allowing early visual processing to be task modulated. Using artificial neural networks, we show that significant improvements in task accuracy can be obtained by allowing the weights to be modulated by task. The primary benefits are obtained under resource-limited processing. A relatively modest task-based modulation of weights and activities can lead to a large performance boost, suggesting an efficient means of increasing effective cortical capacity.

Keywords: neural networks, top-down processing, visual perception, control

Introduction

Individuals typically perceive the world around them with a task or goal in mind. Despite the active nature of perception, the traditional theoretical perspective, exemplified by the early work of Marr (1982), casts the visual system as a static, passive sensory structure that constructs a veridical representation of all facets of the environment, regardless of immediate goals. The neural architecture that embodies this perspective is a rigid hierarchy of feature detectors, which constructs a representation of the visual scene that can be used by subsequent decision making and action systems.

An alternative perspective is beginning to emerge that characterizes visual information processing as dynamic, flexible, and specialized to current goals. According to this perspective, these goals modulate the nature of the visual analysis that is performed, the flow of information within the visual system, and the resulting representations that are constructed.
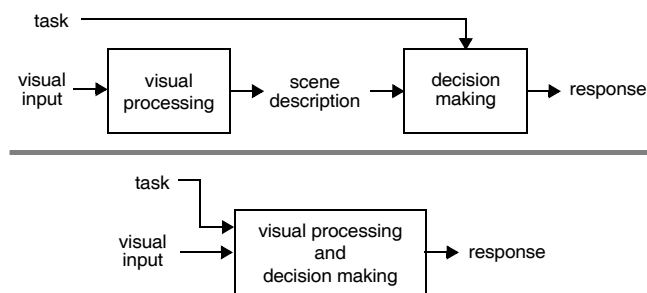

Figure 1: Two visual processing architectures: (a) bottom-up architecture in which visual processing is task independent; (b) top-down architecture in which visual processing is task dependent.

These two contrasting perspectives are depicted in Figure 1. We use the terms *bottom up* and *top down* to refer to the perspectives in which dynamics of visual processing is task independent or task dependent, respectively. Bottom up simply implies that processing is guided from the outside world, and top down implies that processing is guided from higher cortical regions. When we refer to 'bottom up' and 'top down', refer only to processing dynamics, and not to learning, which surely must have a top-down component; and we do not deny that goal-independent processing might well exploit some type of top-down feedback.

The vast majority of cognitive neuroscience models in vision adopt an essentially bottom-up perspective. Although an exciting recent development has been to consider the role of top-down processing (e.g., Bar et al., 2006), "top down" refers primarily to knowledge, expectations, and temporal and spatial context.

In this work, we focus specifically on top-down influence of current goals and the task being performed, and we use "top down" as a synonym for task or goal dependent. We investigate the continuum of possible models by varying the strength of top-down modulation of visual processing based on current goals and tasks. Surely, specializing an architecture for a task at hand should facilitate processing—i.e., it should yield more accurate results given a fixed amount of neural hardware, or it should require less neural hardware to achieve a given level of accuracy. Our exploratory study goes further and investigates the following issues: the magnitude of the benefit of top-down modulation, the conditions under which a benefit is obtained, and the types of top-down modulations that yield benefits. Our research aims to test the viability of the hypothesis that top-down cortical feedback serves to tune the perceptual system to the task at hand.

Evidence for top-down modulation

In this section, we review evidence from neuroscience, neuroimaging, and psychology concerning the top-down modulation of visual processing.

Tasks and goals can influence processing in multiple ways. On a slow time scale, practice on a task leads to learning of new representations. For example, Lee et al. (2002) recorded neural activity while a monkey performed visual search with shape-from-shading stimuli. V1 and V2 neurons show pop-out response to targets that increased with experience and skill. On a somewhat shorter time scale, Zemel et al. (2002) demonstrated that human participants showed an immediate reorganization of perceptual grouping after being exposed to

a set of novel shapes. These studies can be explained by long-term perceptual learning, rather than a dynamic change in neural response properties when the task changes.

A dynamic case, occurring on a time scale of seconds, is the influence of selective attention in the ventral pathway. Attention to a stimulus attribute—location, orientation, direction of motion, etc.—raises the gain of neural responses (Kastner & Ungerleider, 2000). In V4, 25% of neurons show a statistically reliable difference in response depending on a target orientation (Maunsell et al., 1991).

Attentional modulation primarily yields greater sensitivity without much change in selectivity (Maunsell, 2004). Some support has been documented for changes in selectivity. Using binocular rivalry displays, Logothetis and Schall (1989) found neurons in the superior temporal sulcus that reflected the monkey's perception, not the retinal stimulation. Crist, Li, and Gilbert (2001) found an influence of contextual stimuli placed outside the classical receptive field of V1 cells that was consistent with a trained discrimination, but only when the monkey was performing the task.

Evidence for task-based modulation is also found in humans. PET studies show that the same retinal input can activate different extrastriate areas, or even the dorsal versus the ventral stream, depending on whether the task being performed is detection or discrimination (Fias et al., 2002; Orban et al., 1996). In a behavioral study, Schyns and Oliva (1991) found that the perceptual features extracted from images of faces depend on their diagnosticity for the task at hand: With hybrid stimuli combining images at two different spatial scales, a primary discrimination task (gender, expression, or identity) determines which scale is used to make secondary judgments.

The high-density feedback projections found in each layer of the visual hierarchy are naturally interpreted as providing a top-down signal. Beyond the projections to carry the signal, top-down modulation requires the ability to influence cell response properties. Steriade (2004) has argued that slight changes to membrane potential can have dramatic effects: firing patterns can transform a neuron from one response type (regular spiking, fast spiking, fast rhythmic bursting, intrinsically bursting) to another.

Modeling top-down modulation of visual processing

Having summarized the neuroscientific evidence indicating that tasks and goals influence visual information processing, we now discuss what we mean by "influence" in terms of computational models. We adopt the artificial neural network (ANN) paradigm, because ANNs are arguably closer to the underlying neurobiology than are other machine learning models. We acknowledge that ANNs are still quite an abstraction from true neural architectures and dynamics, but past research has been successful in using the ANN paradigm to draw inferences and make predictions concerning the operation of biological systems (e.g., Zipser & Anderson, 1988).

To provide a concrete framework for addressing task-based vision, we focus on visual search: determining the presence
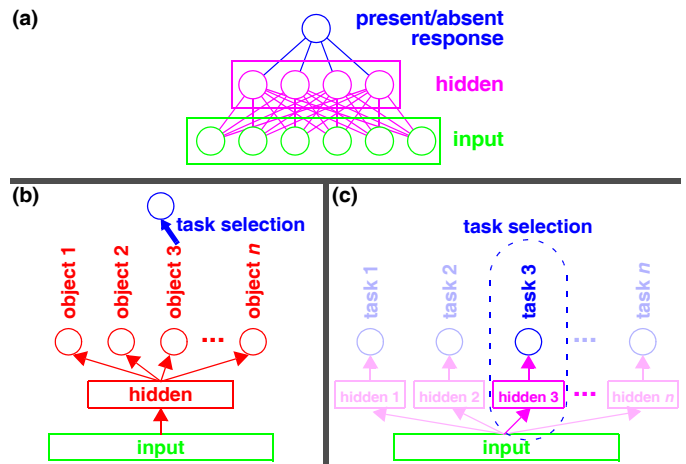

Figure 2: (a) a generic single-task neural network architecture; (b) a bottom-up multitask architecture; (c) a top-down multitask architecture

or absence of a target in a single-item display. Each different target corresponds to a task. For example, with single-digit displays, one could define ten tasks, corresponding to: "Is a 0 present in the display?", "Is a 1 present in the display?", etc.

Figure 2a depicts a generic neural net to perform a specific task, i.e., answer a specific visual search question. The input represents the visual field, and the output a binary present/absent decision. The hidden layer allows for a flexible task-appropriate re-representation of the input. This simple model is not meant to have any isomorphism to the visual system, but is only meant as a functional architecture that performs the same I/O mapping as the visual system (more on this topic in the Discussion section).

Figure 2b embeds the network of Figure 2a into a pure bottom-up architecture. In this architecture, the input is classified into one of n object categories, and following classification, task selection chooses the appropriate output unit to answer the present/absent question associated with the task. In contrast, the top-down architecture in Figure 2c switches the processing that is performed on an input conditional on the task. In this Figure, the n hidden layers are a shorthand for the n task-specific analyses that could be performed on the input. Essentially Figures 2b and 2c differ in that the bottom-up architecture utilizes a hidden representation—specifically the weights from the input to the hidden layer—that is task independent, whereas the top-down architecture utilizes a different hidden representation for each task, allowing for an extreme sort of task dependence.

The bottom-up and top-down architectures are endpoints of a continuum that represents the *strength* of top-down modulation. We explore the continuum by manipulating the degree to which hidden unit weights and activities can be modulated by task. No modulation leads to task-independent weights and a bottom-up architecture; arbitrary modulation leads to task-dependent weights and a strong top-down architecture.

We use α to denote the parameter that controls the strength of task modulation, where $\alpha = 0$ and $\alpha = 1$ correspond to the bottom-up and top-down architectures, respectively, and α plays the following role. In the bottom-up case, we have a single set of input-to-hidden weights, denoted $\bar{\mathbf{w}}$. In the top-down case, we have a distinct set of weights for each task t, denoted $\hat{\mathbf{w}}_t$. Interpolating between these two cases obtains actual weights used for task t, $\mathbf{w}_t$, defined as:

$$\mathbf{w}_t = (1 - \alpha)\bar{\mathbf{w}} + \alpha\hat{\mathbf{w}}_t, \tag{1}$$

where $\bar{\mathbf{w}} = \sum_u \hat{\mathbf{w}}_u / n$ and n is the number of distinct tasks.

Just as the connectivity of the model is not intended to have any neurobiological reality, neither is α. We are simply designing a model that has the necessary *functional* characteristics: α allows us to manipulate the degree of top-down modulation. Using an ANN model for this exploration allows us to manipulate properties of the model—namely, activation and weights—that at least have a rough correspondence to actual neural hardware.

Research questions

Having proposed a simplistic but concrete architecture to examine the role of top-down modulation of information processing, we turn to questions we hope to answer with this line of research. In general, the questions concern the benefit of top-down modulation. Although the neuroscientific and psychological data mentioned earlier provide clear evidence of top-down modulation, cognitive neuroscience has little computational understanding of the magnitude of the benefit, the conditions under which a benefit is obtained, and what types of top-down modulations yield benefits. Consider the following issues.

1. Suppose the bottom-up architecture (Figure 2b) is given the same number of free parameters (weights) as each of the task-specific subnets of the top-down architecture (Figure 2c). This situation corresponds to the case in which top-down modulation has the ability to arbitrarily rewire the network connectivity. Because the top-down architecture benefits from an *n*-fold expansion in representational and computing power, one would be surprised if it did not win out.

2. Suppose the top-down and bottom-up architectures are matched on total number of free parameters (weights). By matching, we mean that a top-down architecture with *h* hidden units is compared to a bottom-up architecture with roughly *nh* hidden units. Two early ANN studies offer conflicting predictions in this situation. Caruana (1997) studied multitask learning in neural nets and found a benefit for sharing of hidden representations among tasks, as would occur with the bottom-up architecture. Rueckl et al. (1989) found that partitioning hidden units to handle specific tasks helped performance. The difference between these two studies was the amount of similarity among tasks. Caruana studied similar tasks, whereas Rueckl et al. studied tasks that appeared to have little overlap with one another.

3. What strength of top-down modulation is necessary to obtain a benefit over the pure bottom-up architecture? That is, for what value of α do we find a benefit? Because the neurobiological evidence suggests a limited modulation due to top-down influences, one would have to question whether top-down modulation played the role suggested by our framework if a benefit is obtained only for α values close to 1. In contrast, if an intermediate strength of top-down modulation (i.e., $0 < \alpha < 1$) outperforms the $\alpha = 0$ and $\alpha = 1$ cases, it would provide a computational rationale for the sort of weak top-down modulation observed in the brain.

4. How does the strength of top-down modulation (α) interact with: (a) the amount of processing resources (hidden units) available, (b) the noise level in the input, and (c) the amount of training data. All three of these variables influence the difficulty of performing each task, and in these performance-limited cases, the benefit of top-down modulation may be more apparent. Such a finding would be consistent with the neuroscientific finding that attentional modulations are larger when displays are cluttered or noisy (e.g., Maunsell, 2004).

Methodology

We studied four data sets, summarized in Table 1. Sets A, C, and D are from the UCI Machine Learning Repository (`pendigits`, `letter-recognition`, and `isolet`, respectively), and Set B is derived from the MNIST database and available at `yann.lecun.com/exdb/mnist/index.html`. Sets A-C are visual images, utilizing three different feature representations: stroke based, pixel based, and statistical moments and edge counts. Set D uses the 26 spoken letters of the alphabet. Because our modeling is at an abstract level, it did not seem unreasonable to include a non-visual data set. Our primary goal in picking data sets was to find a diverse collection using different representations.

For any α, the training procedure involved a search in the model's underlying parameter space, $\{\hat{\mathbf{w}}_t\}$. These parameters are translated into weights used in the activation dynamics, $\{\mathbf{w}_t\}$, via Equation 1. The logic of this procedure is to obtain weights whose strength of task-specific modulation is related to α. The training procedure and the $\{\hat{\mathbf{w}}_t\}$ are not meant to have any neurobiological reality; they simply provide a means of obtaining sets of weights with given functional properties.

Training was performed via on-line steepest descent in the underlying parameter space, $\{\hat{\mathbf{w}}_t\}$, using a mean-squared-error objective function. For each training example on each epoch, gradients were computed with respect to each task's underlying parameters, and weights were updated. An adaptive learning rate was used, which increased by a small constant if the error dropped from one epoch to the next, and decreased by a constant of proportionality if the error rose. To ensure that the network was trained to a local minimum, we used a conservative criterion that terminated training only

Table 1: Data sets used in simulations

label	description	source	input dim	# tasks	# examples	evaluation	# hid
A	Pen-Based Handwritten Digits	UCI	256	10	10992	70% train / 30% test	2
B	Pixel-Based Handwritten Digits	MNIST	196	10	5000	5-fold cross validation	2
C	Distorted-Font Letters	UCI	16	26	20000	3-fold cross validation	15
D	ISOLET Spoken Letters	UCI	617	26	7797	80% train / 20% test	4

when the epochwise mean squared-error dropped by less than 1% over 100 epochs. For each value of α, each hidden-layer size, and each train/test split, we performed 3-8 replications of training with different random weight initializations. (The number of replications was inversely related to the number of data splits, ensuring a roughly fixed number of total runs.)

For each training run, an ROC curve was computed for the test/validation set, and the area under the ROC curve (*AROC*) was determined. AROC is a measure of a model's intrinsic ability to perform a two-way discrimination—present versus absent in the case of our tasks. The AROC measure ranges from 0.5, indicating chance discrimination, to 1.0, indicating perfect discrimination. An AROC score is obtained for each task, and we compute the mean AROC score across tasks.

Results and Discussion

Figure 3 shows results for data sets A-D, with one row per data set. The first column of the Figure plots AROC discrimination performance as a function of the size of the hidden layer, both for the pure bottom-up net ($\alpha = 0$, dashed line) and the strong top-down net ($\alpha = 1$, solid line). The error bars indicate +/- 1 standard error of the mean. AROC discrimination performance improves as the network size increases, at least for the range of network sizes we studied. The x-axis of the graph indicates both the hidden layer size of the bottom-up net, and the hidden layer size for each task in the top-down net. Thus, comparing the bottom-up and top-down AROC values for a given value on the x-axis allows a factor of *n* more free parameters for the top-down net. This assumption is sensible if top-down modulation can completely rewire the available hidden units for each task. At this extreme, it's not surprising that the top-down net outperforms the bottom-up net.

At the other extreme, when no rewiring can take place, the natural comparison is between nets matched on total hardware: a top-down net with *h* units per task versus a bottom-up net with *nh* units. The first column of Figure 3 implies—assuming that performance remains at asymptote as the number of hidden units is increased—little or no advantage for the top-down net in this case. Thus, the benefit of task-specific modulation of unit responses arises in conditions where hardware resources are limited and can be effectively reused by top-down modulation. For this reason, in all subsequent simulations, we selected a particular hidden layer size for each data set, such that resource constraints arose. The size of the hidden layer for each data set is shown in Table 1.
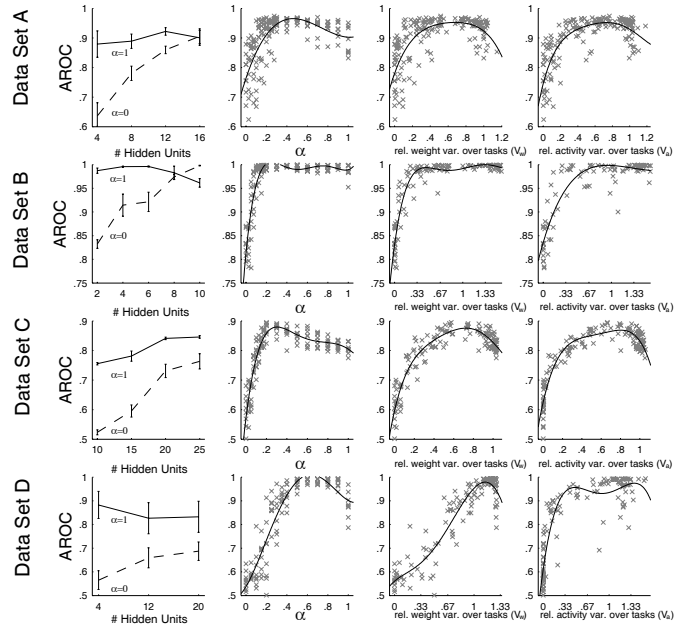

Figure 3: AROC performance on data sets A-D. The graphs are explained in detail in the text. In the first column of graphs, the number of hidden units is varied. In columns 2-4, the number of hidden units is fixed, as described in the text.

The second column of Figure 3 plots the AROC as a function of training α. Individual runs—each with a different weight initialization and/or train/test split—are marked with an "x". The data points are fit with a fourth or fifth order polynomial to give a sense of the relationship. As the strength of top-down modulation of the weights (α) increases, so does performance. Interestingly, in three of the four data sets, an intermediate strength of top-down modulation, e.g., $\alpha = 0.5$ yields reliably better performance than independence of the weights across tasks, $\alpha = 1$ (A: $t(14) = 4.15$, $p < .001$; B: $t(14) = 1.35$, $p > .10$; C: $t(14) = 4.62$, $p < .001$; D: $t(14) = 4.066$, $p < .002$). And importantly, even weak top-down modulation obtains performance improvements. We interpret these results as strong support for the computational benefits of task-dependent modulation of unit responses.

We have treated α as if it controls the degree to which hidden weights and responses are allowed to vary from one task to the next. However, α is primarily a parameter of training, because for $\alpha > 0$, any arbitrary set of weights, $\{\mathbf{w}_t\}$, can be obtained via Equation 1 and appropriate selection of $\{\hat{\mathbf{w}}_t\}$.

Thus, it is necessary to assess the strength of task modulation in other ways.

One measure of task-specific modulation is how much individual hidden-unit weights vary from one task to another. For some hidden unit i having weights $\mathbf{w}_{ti}$ for task t, the variance is $E_t[|\mathbf{w}_{ti} - E_{t'}[\mathbf{w}_{t'i}]|^2]$. We compute the expected variance over weights, and normalize this quantity with respect to the variance within a task, across hidden units:

$$V_w = E_{i,t}[|\mathbf{w}_{ti} - E_{t'}[\mathbf{w}_{t'i}]|^2]/E_{t,i}[|\mathbf{w}_{ti} - E_{i'}[\mathbf{w}_{t'i'}]|^2]. \quad (2)$$

If $V_w = 0$, we have a pure bottom-up net in which the weights do not vary across tasks. If $V_w = 1$, a particular weight varies as much from one task to the next as the weights for a particular task vary from one connection to another. The third column of Figure 3 plots the AROC value of each run as a function of its corresponding V_w value. For data sets A, B, and C, the best AROC value is obtained for V_w in the neighborhood of 0.5, 0.33, and 0.5, respectively. Thus, to achieve optimal performance, top-down modulation must vary the weights from task to task, but the amount of adjustment is far less than the variation one observes over different units.

A second measure of task-specific modulation is analogous to the first, but uses relative *activation* variance instead of weight variance. The measure V_a, depicted in the fourth column of Figure 3, is the expected variance in the activity of a particular hidden unit across tasks, relative to the expected variance in a particular tasks across all hidden units. When $V_a = 0$, we have a pure bottom-up net which yields no variance in the hidden activations across tasks. When $V_a = 1$, a particular hidden unit's activity varies as much across tasks as activity varies for a particular task across units. Similar to the V_w graphs, we find that a a relatively small task-dependent modulation yields significant improvements in AROC; V_a values in the neighborhood of 0.5 yield optimal performance.

The α, V_w, and V_a graphs all point to the same conclusion: *any* degree of top-down modulation of responses yields a significantly increased discriminative ability of the ANN. Although this conclusion is hardly startling, what is surprising and interesting is (a) the slope of the curves, i.e., how sharply performance improves with even small modulations, and (b) the magnitude of improvement that can be obtained by top-down modulation. The slopes on data set D are somewhat shallower than on the other three data sets. This slope cannot be attributed to the number of tasks, because both C and D involve 26 tasks. It might be attributed to the high input dimensionality of set D, which allows any degree of task modulation to affect a large number of weights, consistent with the finding that Set D's V_w curve is shallower than its V_a curve.

We performed additional simulations exploring constrained top-down influences of hidden unit activity involving only the gain or bias of a unit. That is, the net input to hidden unit j is defined as $net_j = g_t(\sum_i w_{ji}x_i) + b_t$, where $\mathbf{x}$ is the input vector, b_t is a task-specific bias, and g_t is a task-specific gain. We allowed either b_t or g_t to vary with task during training, providing additive or multiplicative contri-
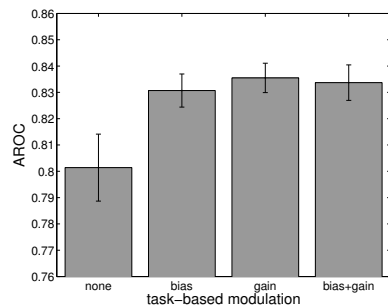

Figure 4: AROC performance on task B for different types of constrained top-down modulation. Modulation of biases, gains, or biases+gains led to only moderate improvements in performance. Error bars show one standard error of the mean.

butions to the otherwise task-independent net input to a hidden unit. These constrained modulations correspond to a subclass of responses observed in neural activity (e.g., Maunsell, 2004).

We found modest benefits for top-down modulation of biases and gains. The largest effect was in data set A, where AROC rose from .76 to .83 ($t(28) = 1.99$, $p < .05$). Effects of bias and gain modulation on data set B are shown in Figure 4. The improvements are small relative to what we observed by allowing task modulation of individual connections. In biological neurons, bias and gain changes are observed in the context of priming phenomena and attentional phenomena, respectively. We speculate that the reason our simulations showed little benefit of these modulations is because our tasks did not involve sequential effects—in which case priming is useful—or multiple-object displays—in which case attentional selection is useful.

In a final set of simulations, we explored the interaction between top-down modulation and input noise. During training and testing, we added Gaussian noise to each input unit i with mean zero, and standard deviation $\lambda\sigma_i$, where λ is a noise level, and σ_i is the standard deviation of the input activity in the training set. Figure 5 shows the outcome for $\lambda \in 0, 0.25, 0.5, 1$ and data sets A and B. Noise had little impact on $\alpha = 0$, most likely because the weights were constrained by all tasks in parallel, and the fewer degrees of freedom led to less overfitting of the training data. Nonetheless, the benefits of top-down modulation over a pure bottom-up model net are obtained for most levels of noise. Even in data set A, where noise had a big impact, top-down modulation leads to a sort of noise robustness: its AROC value for $\lambda = 0$ and $\alpha = 0$ is comparable to that for $\lambda = 0.5$ and $\alpha = 0.2$.

Conclusions and Future Work

Our simulations show a clear benefit of top-down task-based modulation of neural responses. Although the architectures we studied had little of the structure of the human visual system, and the ANN is a highly idealized neural network, we are confident that the results apply to neurobiological sys-
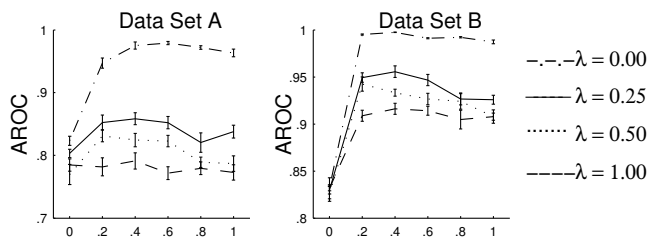
Figure 5: Performance on data sets A and B as a function of α for various noise levels.

tems. If anything, our results are probably conservative because the visual system is a multilayered hierarchy, and each layer introduces nonlinearities. As a result, small modulations to the response of neurons in an early layer can amplify as they propagate forward.

Top-down modulation increases the effective resources of a neural architecture, and is therefore most useful to overcome resource limitations. However, our simulations showed that top-down modulation may also be useful in overcoming some amount of input noise. We observed that a moderate top-down modulation (e.g., $\alpha = 0.5$) yielded better performance than a pure bottom-up ($\alpha = 0$) or a strong top-down model ($\alpha = 1$). In a sense this result should not be terribly surprising. The bottom-up model is ideal if the tasks share strong similarities, and therefore internal representations of the input should be identical. The strong top-down model is ideal if the tasks are unrelated, allowing for flexible representations from task to task. Because the best performance in the tasks we studied was obtained for an intermediate strength of top-down modulation, we can infer that the tasks we studied have an intermediate degree of similarity.

A surprising and interesting finding of our simulations is that even small task-based modulations yield significant performance improvements, as evidenced by the steep slopes of our AROC curves for the α, V_w, and V_a graphs. Because similar tasks should utilize similar representations, it makes sense that not much modulation is required to achieve a performance boost. Our simulations provide a computational justification for observed task-dependent modulation of neural activity, and suggest that experimental studies should be even more sensitive to such modulations.

Our simulations are limited by the number of data sets we explored. We also did not have the opportunity to explore another possible benefit of top-down modulation: the ability to generalize to novel tasks (Caruana, 1997). For example, suppose one was asked to search for a target defined by a combination of color and shape. Previous experience with searching for the color and/or the shape may facilitate search for the novel combination via task-based modulation. Regardless of how such an exploration might turn out, we have demonstrated how task-based modulation can increase the effective processing and representational capacity of a hardware limited system like the brain.

References

Bar, M., Kassam, K. S., Ghuman, A. S., et al. (2006). Top-down facilitation of visual recognition. *PNAS*, *103*, 449–454.

Caruana, R. (1997). Multitask learning. *Machine Learning*, *28*, 41–75.

Crist, R. E., Li, W., & Gilbert, C. D. (2001). Learning to see: experience and attention in primary visual cortex. *Nature Neuroscience*, *4*, 519–525.

Fias, W., Dupont, P., Reynvoet, B., & Orban, G. A. (2002). The quantitative nature of a visual task differentiates between ventral and dorsal stream. *Journal of Cognitive Neuroscience*, *14*, 646–658.

Kastner, S., & Ungerleider, L. G. (2000). Mechanisms of visual attention in the human cortex. *Annual Rev. Neurosci.*, *23*, 315–341.

Lee, T. S., Yang, C. F., Romero, R., & Mumford (2002). Neural activity in early visual cortex reflects behavioral experience and higher-order perceptual saliency. *Nature Neuroscience*, *5*, 589–597.

Logothetis, N. K., & Schall, J. D. (1989). Neuronal correlates of subjective visual perception. *Science*, *245*, 761–767.

Marr, D. (1982). *Vision*. San Francisco, CA: W. H. Freeman.

Maunsell, J.H.R. (2004) The role of attention in visual cerebral cortex. In L.M. Chalupa and J.S. Werner (Eds.), *The visual neurosciences* (pp. 1538–1545). Cambridge, MA: MIT Press.

Maunsell, J.H.R., Sclar, G., Nealey, T.A., DePriest, D.D. (1991) Extraretinal representations in area V4 in the macaque monkey. *Visual Neuroscience*, *7*, 561–573.

Orban, G.A., Dupont, P., Vogels, R., De Bruyn, B., Bormans, G., & Mortelmans, L. (1996). Task dependency of visual processing in the human visual system. *Behav. Brain Res.*, *76*, 215–223.

Rueckl. J.G., Cave, K.R. & Kosslyn, S.M. (1989). Why are "what" and "where" processed by separate cortical visual systems? A computational investigation. *Journal of Cognitive Neuroscience*, *1*, 171–186.

Schyns, P., G., & Oliva, A. (1999). Dr. Angry and Mr. Smile: When categorization flexibly modifies the perception of faces in rapid visual presentations. *Cognition*, *69*, 243–265.

Steriade, M. (2004). Neocortical cell classes are flexible entities. *Nature Reviews Neuroscience*, *5*, 121–134.

Zemel, R. S., Behrmann, M., Mozer, M. C., & Bavelier, D. (2002). Experience-dependent perceptual grouping and object-based attention. *Journal of Experimental Psychology: Human Perception & Performance*, *28*, 202–217.

Zipser, D., & Andersen, R. A. (1988). A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons. *Nature (London)*, *331*, 679–684.