StatTexNet: Evaluating the Importance of Statistical Parameters for Pyramid-Based Texture and Peripheral Vision Models

C. Koevesdi
V. DuTell
A. Harrington
M. Hamilton
W. T. Freeman
R. Rosenholtz

MIT CSAIL, Brain and Cognitive Sciences

Abstract

Peripheral vision plays an important role in human vision, directing where and when to make saccades. Although human behavior in the periphery is well-predicted by pyramid-based texture models, these approaches rely on hand-picked image statistics that are still insufficient to capture a wide variety of textures. To develop a more principled approach to statistic selection for texture-based models of peripheral vision, we develop a self-supervised machine learning model to determine what set of statistics are most important for representing texture. Our model, which we call StatTexNet, uses contrastive learning to take a large set of statistics and compress them to a smaller set that best represents texture families. We validate our method using depleted texture images where the constituent statistics are already known. We then use StatTexNet to determine the most and least important statistics for natural (non-depleted) texture images using weight interpretability metrics, finding these to be consistent with previous psychophysical studies. Finally, we demonstrate that textures are most effectively synthesized with the statistics identified as important; we see noticeable deterioration when excluding the most important statistics, but minimal effects when excluding least important. Overall, we develop a machine learning method of selecting statistics that can be used to create better peripheral vision models. With these better models, we can more effectively understand the effects of peripheral vision in human gaze.

Keywords: peripheral vision, texture synthesis, multi-scale pyramid, statistic selection, contrastive learning

1. Introduction

A key source of information in human gaze comes from peripheral vision. While it is often thought of as an adaptation to capacity limits of the human visual system, peripheral vision also drives human performance on many visual tasks – including search, scene perception, and object detection (Ehinger and Rosenholtz, 2016). With respect to gaze specifically, peripheral vision plays a role in saccadic planning by helping humans determine where to look next (Schütz et al., 2011).

Given its importance in understanding human gaze patterns, numerous attempts have been made to model peripheral vision. Multi-scale-pyramid-based models are the current
state of the art. These models account for both the loss of photoreceptor density and the summarization of information thought to occur in brain areas V2 and V3. Models such as these treat peripheral vision as a texture-like representation and have a long history in human vision. They have been used in not only peripheral vision, but also in texture models more generally (Portilla and Simoncelli, 2000). To simulate peripheral vision, these models utilize overlapping pooling regions that encircle the fovea and increase in size with eccentricity. While some models utilize machine learning techniques like style transfer (Wallis et al., 2017; Deza et al., 2017) to summarize information, the majority of these models calculate summary statistics for each pooling region, which are calculated on the output of multi-scale pyramids.

One challenge of pyramid-based models of peripheral vision is in determining which statistics are calculated in each pooling region. Although most pyramid-based texture models used to study peripheral vision have been validated through human behavioral studies, they still utilize statistic sets that are historically driven, vary study-to-study from previous literature, and are consistently insufficient to capture the wide variety of possible textures Brown et al. (2021).

The problem of selecting which statistics are necessary and sufficient to represent the variety of textures perceived in peripheral vision is critical for the goal of building better models of human gaze. While testing every single texture by hand or with a human-in-the-loop is not feasible, we leverage self-supervised approaches in machine learning to address the problem of statistic selection in peripheral vision models. In this work, we develop a constrastive learning model, StatTexNet, to explore the relative importance of pyramid-based statistics for representing peripheral vision. To validate our machine learning approach to statistic selection, we test our framework on a set of depleted texture images with known statistics. We demonstrate that StatTexNet selects the known most important statistics in these depleted textures. We then apply our model to full texture images and use weight interpretability metrics to determine what are the most important statistics to represent texture families. Finally, we synthesize textures using statistics selected by our method.

By building a machine-learning-driven approach to statistic selection, our work automates the evaluation of statistics used by texture-based peripheral vision models. With a better method of understanding and evaluating peripheral vision models, we can build a more complete understanding of human gaze.

2. Previous Work

Peripheral vision represents the majority of the visual field, and both critically limits and enables human performance at a variety of tasks (Rosenholtz, 2016). This includes gaze behavior where information from both the fovea and the periphery are integrated to inform saccades (Stewart et al., 2020).

Some of the best performing models of peripheral vision use a multi-scale pyramid approach. Most pyramid-based peripheral vision models are based on work from the texture modeling world. Early work in this area included (Julesz, 1962), who first explored different textures that could be represented as the same N-th order pixel statistics. Large improvements were seen with a move from pixel-based to multi-scale pyramid based sta-
tistical representations (Simoncelli and Freeman, 1995). The steerable pyramid has since been widely used in vision modeling as its filters resemble those found in the mammalian early visual system (Turner, 1986; Malik and Perona, 1990), which break down an input image into distinct spatial frequency and orientation bands. Using the steerable pyramid, Heeger and Bergen (Heeger and Bergen, 1995) proposed a statistics set calculated on this pyramid decomposition, alongside a histogram-matching procedure that enabled good texture synthesis. This was refined further by (Portilla and Simoncelli, 2000), which included pixel, autocorrelation, and magnitude statistics.

When these texture models were first applied to peripheral vision (Rosenholtz et al., 2012; Freeman and Simoncelli, 2011), they utilized a similar texture set to (Portilla and Simoncelli, 2000). Statistics were modified from this set by being hand-chosen and tested for necessity and sufficiency through trial and error on a limited test set of textures. More recent work has modified these statistics slightly, tested them on a wider variety of conditions and textures, and made code more flexible and efficient (Brown et al., 2021; Wallis et al., 2017).

Behavioral evidence supports the statistics set used by these state-of-the-art peripheral vision models. These models are often used to create mongrels, also known as metamers, which are visual stimuli that match another in representational space, but can differ significantly in pixel space. When viewed foveally, the pixel-differences are obvious, but when viewed peripherally, they are indistinguishable. Mongrels have been shown through careful psychophysical experimentation to reproduce the same capabilities and limitations of human peripheral vision including crowding (Balas et al., 2009) and scene perception (Ehinger and Rosenholtz, 2016). In addition, the scaling parameters for pooling regions needed to create metamers/mongrels mirror those of neuron receptive fields in non-human primates (Freeman and Simoncelli, 2011).

Despite the success of these models, it is clear that the current state-of-the-art statistic set is insufficient. A faithful model of human peripheral vision should work regardless of input type. However, investigations into the effect of different texture families have revealed that for current models, textures with certain properties are more faithfully represented, while metamers/mongrels of other texture types consistently fail (Brown et al., 2021; Broderick et al., 2023). These problems occur despite modifications to optimization strategy, hyperparameters, and seed.

Some efforts have worked to eliminate the need to choose specific statistics altogether. Mongrels have been successfully created by taking inspiration from style transfer (Gatys et al., 2016), utilizing the entire gram matrix as the statistical representation to create metameric images (Deza et al., 2017; Wallis et al., 2017). While this removes the need for hand-picking statistics, this represents a huge matrix that is likely over-parameterized, and removes any potential compression advantage. Another example is the work from (Serre et al., 2007), which simply takes the maximum output of each pooled area. Although the field has made significant progress toward improving the statistic component of peripheral vision models, it is clear that a more principled approach to selecting the statistics is needed.
3. Modeling Textures Through Statistics

In order to build a better method of selecting the most important statistics for texture-based peripheral vision models, we devise a contrastive learning model, StatTexNet, to take a large set of statistics and compress it to a smaller set. In our model, we take 5-crops (4 corners and center) from a texture image dataset, and calculate their summary statistic representation using the GPU-optimized code from (Brown et al., 2021) (Figure 1). This consists of convolution of each 128x128 pixel crop with a steerable pyramid filter bank, and the calculation of 150 summary statistics from these pyramid images. We then use a single fully connected layer to compress this statistical representation, which we train through contrastive learning. The input space is thus 150. For the output latent space, we choose a dimensionality of 50, as it provides the most effective clustering in our experiments.

While we use the statistics set from the (Brown et al., 2021) model as a baseline, we note that this is a similar statistics set to other popular models (Portilla and Simoncelli, 2000; Freeman and Simoncelli, 2011; Rosenholtz et al., 2012), with some statistics removed for computational savings, simplicity, and based on empirical findings, as well as the inclusion of an additional statistic set, ’end-stopped’.

4. Summary Statistics Sets

StatTexNet starts with an initial set of summary statistics which are are split into two groups: first-order and more complex second-order and higher statistics.

Following (Brown et al., 2021), we utilize the following statistics:
**First-order statistics:**

- From the raw input image pixels, the first four moments — mean, variance, skewness, and kurtosis — of the grayscale histogram.
- The variance of both the high- and low-pass bands, with skewness and kurtosis also computed for the latter.
- For the non-oriented lowpass bands, the variance, skew and kurtosis are computed.
- For each bandpass filter output, the magnitude-mean and variance are derived.

**Second- or higher-order statistics:**

- Magnitude-correlations between bandpass filters. This involves the correlations between all orientations at the same scale in the steerable pyramid, but also correlations between neighboring scales at the same orientations.
- The same correlations are also computed for the phase images.
- Finally, unique to Brown et al is the *End-Stopped* statistic. This statistic is based on end-stopped neurons or hypercomplex cells in visual cortex ([Hubel and Wiesel, 1959](#)), and differentiates between segmented and continuous lines. Specifically, each edge magnitude component image is subtracted from a slightly shifted version of itself, following the expected edge direction. The resulting difference is then squared.

![Figure 2](image)

**Figure 2:** We generate depleted textures created with a known set of statistics, feeding these controlled images to our model, and perform the same procedure as in Figure 1.

Natural images are essentially unbounded by the set of statistics that represent them. However, synthesized images are created using only a set of known statistics. In order to control for the set of statistics present in a given texture image and validate our method of
statistic selection, we create synthesized versions of each texture using the Heeger & Bergen texture model (Heeger and Bergen, 1995). Heeger and Bergen preforms histogram matching on first order statistics only and thus its syntheses are only constrained to this subset of statistics. These synthesized textures are depleted in that they do not contain the full set of statistics needed to fully describe them. They can therefore be used to validate our method, as a model should not need higher-order statistics to group them. This enables us to test if a network can learn the relative importance of different groups of statistics from different datasets. To do this, we then follow the same pipeline as in Figure 1, with these depleted images (Figure 2).

5. Datasets

![Dataset visualization through sample textures. The top row indicates the original texture and the bottom row shows the synthesized texture through the Heeger and Bergen procedure.](image)

In this study, we utilized two primary datasets: the Describable Textures Dataset (DTD) (Cimpoi et al., 2014) and the KTH-TIPS2-b (KTH) dataset (Mallikarjuna et al., 2006) which we use for validation. The DTD captures a wide array of textures found in natural settings and is a collection of 5,640 images spanning 47 distinct texture categories. These images were primarily sourced from platforms like Flickr and Google Search. The KTH dataset contains 4,752 images representing 11 different materials that were acquired through imaging 4 different samples for each material, each under varying pose, illumination and scale. Due to the way it was collected, DTD has more intra-class variation than KTH.
We transformed all RGB images from these datasets into grayscale. We then applied the Heeger and Bergen texture synthesis procedure (Heeger and Bergen, 1995) to these grayscale images. The Heeger and Bergen approach is to iteratively modify a gaussian white noise image so that the pixel distributions in its steerable pyramid representation match that of the reference texture. This is done through histogram matching. When provided an input image, histogram matching aims to adjust the image’s grayscale pixel value distribution so that it aligns with the histogram of a reference image. Thus, histogram matching adjusts the pixel distribution of an image to match that of a reference, ensuring identical first-order statistics, but not guaranteeing similar spatial structures or correlations between images.

Consequently, we have four datasets at our disposal to test our hypotheses: two are the original grayscale sets (DTD and KTH), and the other two are depleted - derived from the Heeger and Bergen synthesis method applied to DTD and KTH. Figure 3 shows some examples of these datasets.

6. Training

6.1. Contrastive Learning

Our goal is to reduce the full set of 150 Brown et al. (2021) image statistics to a compressed representation 1/3 the size, forcing the network to prioritize information from certain textures over others. To do this, we employ constrastive learning (Chen et al., 2020), allowing our network StatTexNet to learn any representation that is useful in discriminating textures. Contrastive learning works by ensuring that similar pairs, such as crops from the same image, are drawn close together in representation space, while distinct pairs are pushed apart based on a specified distance measure. For this task, we utilize generalized lifted structured loss (Hermans et al., 2017) with a Euclidian distance. The advantage of this loss is its ability to effectively process the entire training batch, taking into consideration both closely related pairs (positive anchors) and those that are unrelated (negative pairs). In one training step, all pairs are considered (See Appendix Section 12.2).

For our input data, we take a single texture image from one of our datasets and crop it into 5 smaller images. This gives use a set of 5 images that we know come from the same texture, and thus, should be represented by a very similar statistical values. Crops from the same image are treated as positive samples and crops from different texture images are treated as negative in our framework. We train our contrastive learning networks for 200 epochs. (See 12.3 for details on data augmentation). To process our data efficiently and ensure consistent gradient updates, we selected a batch size of 100. Additionally, after evaluating different optimization techniques, we settled on the Adam optimizer due to its adaptive learning rate and proven success in similar tasks. We used a learning rate of 0.0001.

6.2. Dropout

One complication of our model is that correlations between different elements of the 150 statistics set could potentially cause the network to ignore certain highly correlated or anti-correlated statistics. We reasoned that, because our statistics sets represent uniform spatial samples of natural images that have inherent regularities (Ruderman, 1997; Simoncelli and Olshausen, 2001), correlation between statistics was highly likely. This could happen when
multiple statistics correlate sufficiently such that the network learns to rely only on one of the correlated statistics, discounting others.

To address this, we first checked for correlations among statistics (Appendix Section 12.5), and found that indeed, the majority of statistics measured show high correlation with at least one other. We counteract this issue by incorporating dropout during training. By incorporating dropout, some features are set to zero temporarily at random during each forward pass. This prevents the model from becoming too reliant on specific features as it forces the model to learn a more even distribution across correlated groups. Thus, this approach mitigates the effects of multi-collinearity. We find that incorporating dropout greatly improves the results in Table 1, compared to training without dropout.

![Figure 4: t-SNE visualizations for learned embeddings of the DTD dataset for both original texture images (left), as well as depleted (right). The network learns to cluster textures from both well, while the right plot indicates that depleted images are clustered better.](image)

### 6.3. t-SNE

In addition to seeing a reduction in loss over training, we validated the effectiveness of our contrastive learning approach using t-SNE (Van der Maaten and Hinton, 2008) to visualize the learned latent representation space. To do this, we ran inference on a set of 20 randomly-chosen textures, each with 5 crops, and visualized the 2D embedding of the 50 dimensional space (Figure 4). We find that indeed, crops from the same texture cluster together well in space. This is especially true for the depleted textures synthesized with (Heeger and Bergen, 1995).

### 7. Rankings for Depleted Data

#### 7.1. Weight-Based Ordering

After the training process, we do weight-based ordering on StatTexNet to determine the significance of each statistic. We summed up the absolute weight values for each input node
统计学网络：评估周围统计量

表1：50个第一阶总结统计量的重要性指标，基于两种不同的特征选择方法。对于两种排序，所有三种度量和两个数据集，第一阶统计量在去除纹理中被创建时，比在原始纹理中的统计量更重要（即排名更高），只有当使用这些统计量时，才比原始纹理。

<table>
<thead>
<tr>
<th>Ordering</th>
<th>Metric</th>
<th>DTD</th>
<th>KTH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Original</td>
<td>Depleted</td>
</tr>
<tr>
<td>Weight</td>
<td>% in Top 15</td>
<td>36.00</td>
<td>75.33 ✓</td>
</tr>
<tr>
<td></td>
<td>Median rank</td>
<td>59.45</td>
<td>41.60 ✓</td>
</tr>
<tr>
<td></td>
<td>Mean rank</td>
<td>63.36</td>
<td>52.65 ✓</td>
</tr>
<tr>
<td>Shapley</td>
<td>% in Top 15</td>
<td>32.67</td>
<td>49.33 ✓</td>
</tr>
<tr>
<td></td>
<td>Median rank</td>
<td>71.20</td>
<td>49.45 ✓</td>
</tr>
<tr>
<td></td>
<td>Mean rank</td>
<td>67.75</td>
<td>55.55 ✓</td>
</tr>
</tbody>
</table>

在我们的模型中，每个输入节点对应一个150个特征统计量之一。因为我们在权重矩阵中使用了缩放来归一化，所以一个更有用的统计量在分类中应该被网络赋予更高的权重。我们按照降序排列这些权重，从最不重要（低排名）到最重要（高排名）。

我们遵循了三个指标来评估StatTexNet是否可以学习到最重要的统计量，以应对不同的数据集。我们假设第一阶统计量，符合Heeger和Bergen合成的，将在去除纹理数据中起更重要的作用，比在原始纹理中的作用。为了测试这个假设，我们为每个数据集计算了权重，并按照重要性排名150个统计量。作为第一个指标，我们观察了前15个最重要的第一阶统计量在整体重要性排名中的比例。我们期望在合成纹理数据集中，这些最重要的统计量的比例会更高。而且，我们还评估了50个第一阶统计量的中位数和平均重要性排名。我们在10个不同的种子中观察了这三个指标的平均值，并在所有情况下观察到了一致的结果。

我们发现，对于两个数据集，以及所有3个度量（%在Top 15，中位数排名，平均排名），相对排名反映了我们预期的结果。也就是说，当StatTexNet在去除数据集上进行训练时，它将50个第一阶统计量的排名提高了（更重要）。而在原始纹理数据集的训练上，则提高了（更重要）。对于KTH数据集，100%的前15个统计量属于第一阶统计量中被去除的纹理数据。这发生在所有10次不同随机种子的单独训练中。这表明我们的框架是有效的，基于权重的排序能够识别对比学习任务中最重要的和最不重要的统计量。

我们发现，对于两个数据集，以及所有3个度量（%在Top 15，中位数排名，平均排名），相对排名反映了我们预期的结果。也就是说，当StatTexNet在去除数据集上进行训练时，它将50个第一阶统计量的排名提高了（更重要）。而在原始纹理数据集的训练上，则提高了（更重要）。对于KTH数据集，100%的前15个统计量属于第一阶统计量中被去除的纹理数据。这发生在所有10次不同随机种子的单独训练中。这表明我们的框架是有效的，基于权重的排序能够识别对比学习任务中最重要的和最不重要的统计量。
7.2. Shapley Value-Based Ordering

While weight-based ordering using the average absolute value of weights offers strong support of our hypothesis that depleted data would favor first-order statistics more heavily, we sought a more sophisticated mathematical approach to test and validate our findings. Calculating Shapley values (Roth, 1988) is an interpretability method based in game theory enabling the assignment of credit to individual inputs for a given output in a machine learning model. We utilized the SHAP package (Lundberg and Lee, 2017) to calculate Shapley values for each of the 150 statistics, and used these values in place of absolute value of weights to order statistics by importance.

We find that the rankings based on Shapley values also support our hypothesis that depleted texture-trained networks will more heavily rely on the 50 first-order statistics than networks trained on their complete texture counterparts (Table 1, Bottom). Given these results indicating the strong utility of ranking via Shapley values, we chose to utilize this ranking procedure alongside weight in exploring the statistical importance for non-depleted data.

8. Statistical Importance for Original Textures

Figure 5: Statistic rankings for two datasets tested. Small points indicate individual statistics, large points indicate group statistic means (circle). Phase-correlation statistics are consistently of low importance, while most other statistics families show heterogeneous performance. Shapley ranking of statistics shows better correlation between datasets tested.
Having validated that our method works in identifying the most and least important statistics for texture representation, we turn to the results on original (non-depleted) textures. First, to understand the relative importance of each statistic, we computed the mean ranking of the nine statistics groups (Figure 5, bar plots in Figure 8), averaged over 10 seeds.

We find that overall, bandpass variance (a single statistic) has high ranking between both datasets and ranking procedures (especially for DTD), indicating that it is important. Magnitude-mean statistics also cluster consistently towards high rankings. Most of the other statistics show a wide distribution of rankings. This is true across datasets, within datasets, and for both ranking systems. End-stop and magnitude-correlation statistics in particular show highly distributed rankings, appearing as both some of the most and least important statistics.

We find that phase-correlation is consistently ranked far lower than all other statistics classes, with strong rank grouping near the end, indicating that it is a less important statistic overall. Our findings of phase-correlation being less important are consistent with previous psychophysical literature Balas (2006), which found phase-correlation to be unimportant for discriminating textures. Interestingly, only our weight-ranked results are consistent with their findings that marginals are highly important for discrimination.

9. Synthesis

One advantage of the texture/peripheral models studied here is the ability to synthesize textures based on a given statistics set. This allows us to visually validate our results. While we emphasize that synthesis results have high variation being both highly seed and texture dependent (Brown et al., 2021; Broderick et al., 2023), we nonetheless include some syntheses here, demonstrating the effect of depleting various statistics.

We show examples of textures with properties found by Brown et al. (2021) to be most and least well-captured by the full texture set. We find that high contrast textures like the lined texture, demonstrate similar performance to baseline (All) when the less-important phase-correlation statistic is removed, but fail completely when the highly-important magnitude-mean statistic is removed. Lower contrast textures, like the painted image, however, show similarly poor synthesis in all cases. The porous texture, lying somewhere in between, has similar synthesis performance to baseline when phase-correlation is removed, and a slightly worse performance when magnitude-mean is removed. Our observations in Figure 8 align with this, highlighting that the magnitude-mean statistics are notably important compared to the phase-correlation statistics. Given that the phase-correlation statistics comprise a greater number of statistics than magnitude-mean, this offers a meaningful point of comparison.

These syntheses support the results uncovered here through our contrastive learning approach. While the 150 statistics of Brown et al. (2021) are not sufficient for all textures, removal of the phase-correlation statistic is often not important, while removal of the magnitude-mean statistic is often noticeable, and sometimes catastrophic.
Figure 6: These three textures represent synthesis failures and success classes based on Brown et al. (2021). High contrast (first row, lined), middle contrast (middle row, porous) and low (bottom row, painted). Low roughness/coarseness textures (bottom row) have poor syntheses for even the full statistics set. Magnitude-mean is important for high and middle contrast textures as shown in the first two rows. Phase-correlation can be removed without much quality loss as compared to the full statistics synthesis.

10. Discussion

In this work, we combine self-supervised learning with weight interpretability analysis to develop, validate, and use a novel method that enables the principled selection and prioritization of the texture summary statistics underlying modern peripheral vision models. By adding a single fully-connected layer to a texture model, we create StatTexNet which we train with contrastive learning to prioritize the most important statistics on the task of grouping textures from the same family together. We show that StatTexNet successfully learns to group textures – indicating that it learned an optimal statistical representation of texture.
In addition, we use multiple weight interpretability metrics to order the relative contribution of individual statistics. To validate this ordering, we create a depleted texture set which is synthesized with a reduced set of statistics, train our network on these textures, and confirm that these reduced set of first-order statistics are the most important in grouping depleted textures as compared to original ones. We show that this result is consistent for 6 different orderings/metrics across 2 different datasets, averaged over multiple seeds.

Finally, we use this method to identify the relative importance of statistics in representing natural textures. When averaging over the sometimes heterogeneic texture families, we find that bandpass variance and magnitude-mean are the most important overall, while phase-correlation is least important. We show that our results are consistent not only with a small sample of synthesized textures, but also with previous psychophysical literature (Balas, 2006), which used psychophysical methodology to evaluate discrimination abilities for depleted textures. While their results found marginal statistics among the most important for the task of texture discrimination, like our work they find that cross-scale phase statistics to be among the least important for this task.

Overall, our method demonstrates a novel, efficient, and principled approach to selecting the statistics for peripheral vision models, as well as the pyramid-based texture-based models that underlie them. While a human in the loop will likely always be necessary to fully validate a statistics set, our method can make such experiments more directed, as testing all possible subsets of even 150 statistics in a formal eye-tracked psychophysics experiment is not feasible.

Future work could scale-up our approach using the larger set of statistics from models such as Portilla and Simoncelli (2000); Freeman and Simoncelli (2011); Rosenholtz et al. (2012), or a novel, much larger set of possible statistics. Additionally, the human visual system is thought to use highly complex transforms and performs a variety of tasks beyond grouping textures. Our method could be utilized to explore the effect of modeling more complex transformations on statistical importance, or the effect of alternative tasks such as classification, as more complex multi-layer weight structures are compatible with the Shapley method demonstrated here. Overall, with our principled and scalable approach to statistic selection, we can work toward better models of texture, peripheral vision, and human gaze as a whole.

11. Acknowledgements

This work was funded by the CSAIL MEnTorEd Opportunities in Research (METEOR) Fellowship, US National Science Foundation under grant number 1955219, as well as National Science Foundation Grant BCS-1826757 to PI Rosenholtz. The authors acknowledge the MIT SuperCloud Reuther et al. (2018) and Lincoln Laboratory Supercomputing Center for providing HPC resources that have contributed to the research results reported within this paper.
References


