

Characterizing the Receptive Field Structure of Visual Neuron Using Supervised and Unsupervised Neural Networks and Pursuit Models

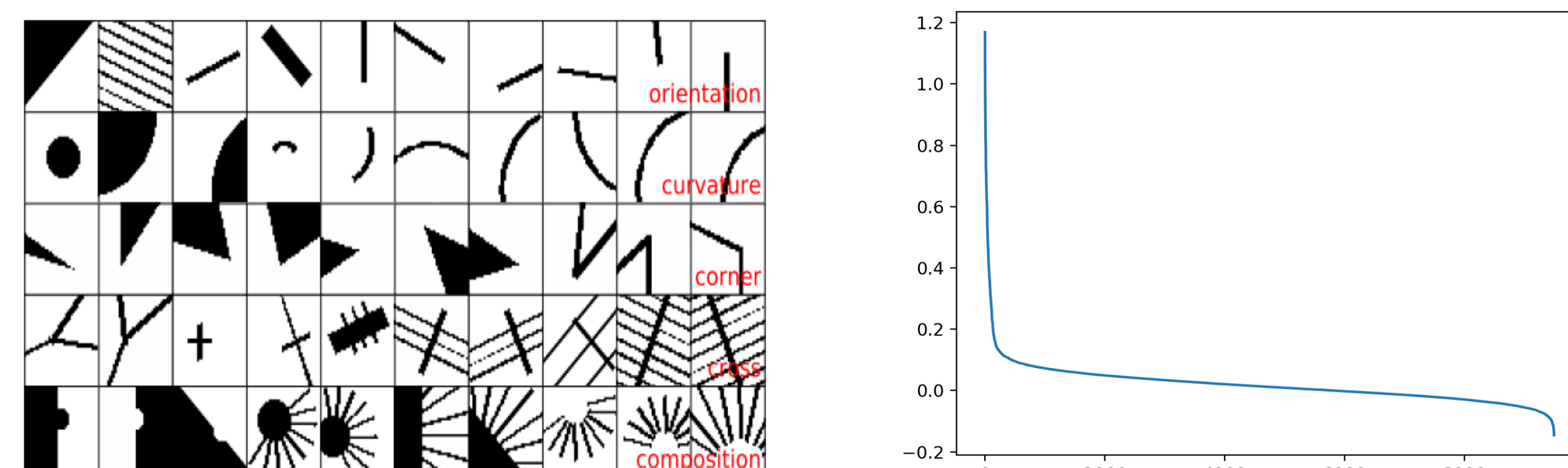
Ziniu Wu, Tai Sing Lee

Goal and Motivation

- This project evaluates and compares the performance of multiples variants of convolutional neural network and pursuit models for modeling the V1 neurons of awake macaque monkeys in response to large set of complex pattern stimulus. **The main purpose is to recover the shape of neurons' receptive field.**
- CNN produce the state-of-the-art results in terms of prediction, but their filters (ie. the recovered RF of neuron) are noisy, lack of **interpretability**. We managed to use **the transfer learning** approach to produce a better characterization of the neuron's RF structure.
- Our conjecture is that there is a set of **more interpretable basis functions that span the same space** (as those spanned by the CNN filters). We tested this conjecture by obtaining a set of complete and over complete dictionary on unsupervised convolutional sparse coding, which produces a dictionary of filters with much less noise. **A better characterization of the neurons' component subunits allow us to have a better way of classifying the neurons, understanding the development of their "compositional" complex selectivity.**
- Traditional approaches fit neuronal response to all stimulus in order to recover neuron's receptive field shape. However, V1 neurons are highly selective to particular patterns making their response to other patterns similar to noise. My model focuses on **top stimulus**.

Dataset and Preprocessing

- Stimulus:** artificial images of complex patterns. 9500x40x40



- Figure 1:** Input pattern images from five classes, Neuron's tuning Curve
- Response:** dF/F roughly has a mean of 0 and standard deviation of 0.1, suggesting that stimulus with response below 0.2 will not be significantly different from a baseline firing rate under 90% CI. Subset with dF/F > 0.2.
 - Resulting dataset: X: input images: 9500x20x20, Y: response 781 neuron.

Finding Interpretable Basis Space

- Patch Based **Sparse Coding (SC)**: $\mathcal{L}(Y, Z; B) = \frac{1}{2} \|Y - BZ\|_2^2 + \lambda \|Z\|_1$
- Convolutional Sparse Coding (CSC):
 $\mathcal{L}(x, z, D) = \frac{1}{2} \|x - \sum_{k=1}^K D_k * z_k\|_2^2 + \lambda \|z\|_1$
 $\mathcal{L}(x, z, D, W) = \frac{1}{2} \|x - \sum_{k=1}^K D_k * z_k\|_2^2 + \sum_{k=1}^K \|z_k - f(W^k * x)\|_2^2 + \lambda \|z\|_1$
- Learned Dictionary (Basis Space):

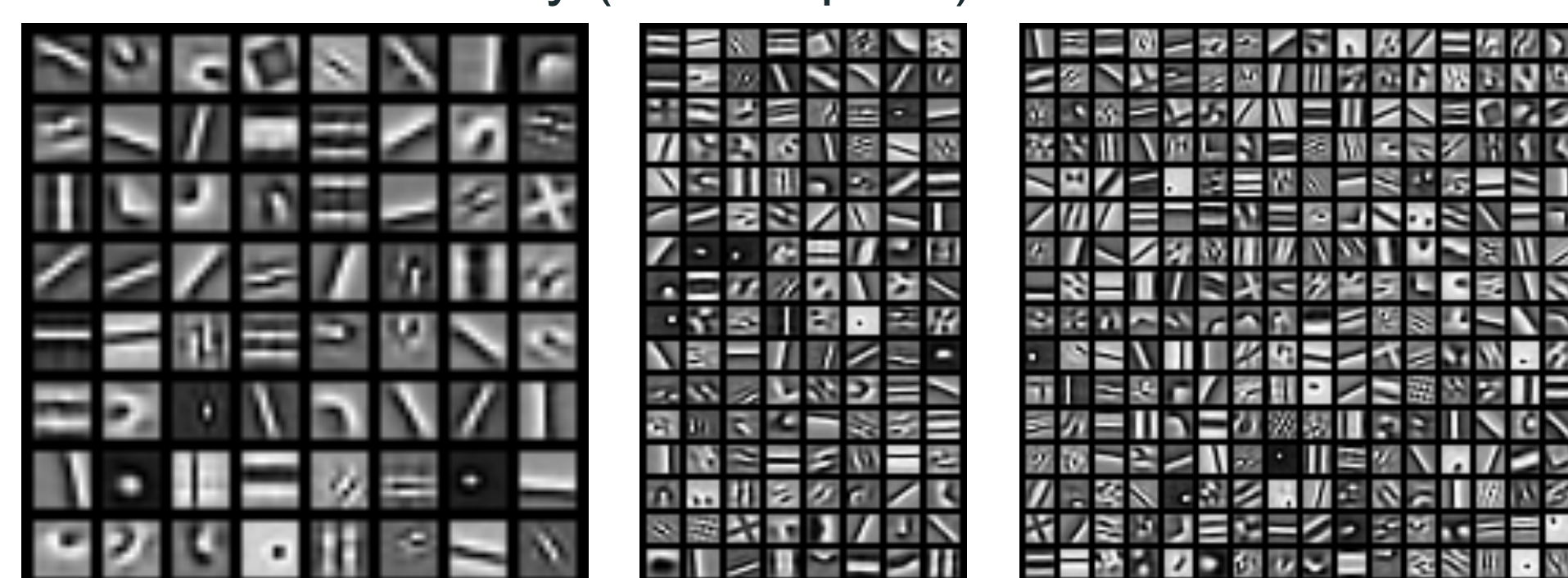


Figure 2: Learned Feature Space with 64, 128, 256 filters, corresponding to different level of completeness

Model and Evaluation

- Projection Pursuit Regression (PPR):** Linear filter, non-linear prediction.
- Convolutional Projection Pursuit Regression (CPPR):** 4 subunits. 13x13 kernel
- Convolutional Matching Pursuit Regression (CMPR):** Using CPPR to select filters from the learned dictionary.
- Semi-matching Pursuit Regression (SMPR):** Select first layer and learn consecutive layers.
- Convolutional Neural Network (CNN):** State of model
- Fixed Kernel Convolutional Neural Network (SCNN):** CNN with kernel fixed by the learned feature space. Neurons will have the same joint filter space but different weight. Neuron's RF structure can be visualized by seeing the filters that has top weights.
- Evaluation: 5-fold cross validation and concatenate the testing result to compute one overall Pearson correlation.
- Compare CMPR with CPPR and SMPR, compare SCNN with CNN, controlling the number of parameters. (Model under comparison with have roughly same number of parameters.)
- Cell classification: classify visual neuron based on the shape of selected filter.

Algorithm 1 CPPR Algorithm

- 1: procedure CPPR(X, y)
- ▷ input set of images and corresponding neural response
- 3: $\psi(X) = \sum_{m=1}^M \text{plog}(F_m * X, 2)$
- 4: $r_i \leftarrow y_i$
- 5: $m \leftarrow 0$
- 6: while $m \leq M$ do
- ▷ M: Number of filters expected
- 7: $I(F) = 1 - \sum_{i=1}^n (r_i - \text{plog}(F * X, 2))^2 / \sum_{i=1}^n r_i^2$
- 8: $F_m \leftarrow \text{argmax}(I(F))$
- 9: $r_i \leftarrow r_i - \text{plog}(F_m * X, 2)$
- 10: $m \leftarrow m + 1$

Figure 3: Projection Pursuit Algorithm

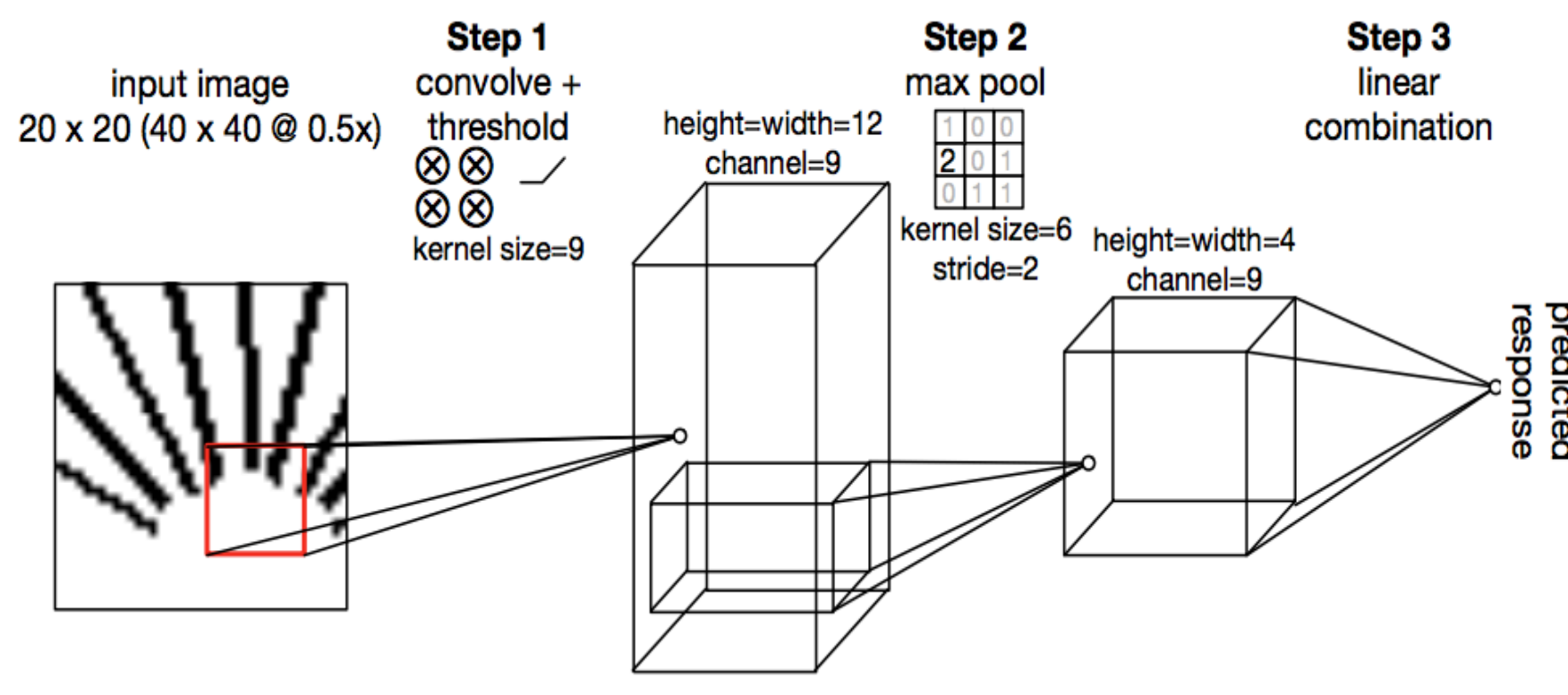


Figure 4: CNN architecture

Characterization

- We will compare this method's characterization of neuron's RF structure with our previous paper's method of classifying neuron.
- Previous method: Manually see the stimulus that produce neuron response larger than 0.5 (roughly 10 stimulus) and classify this neuron based on the shape of these stimulus.

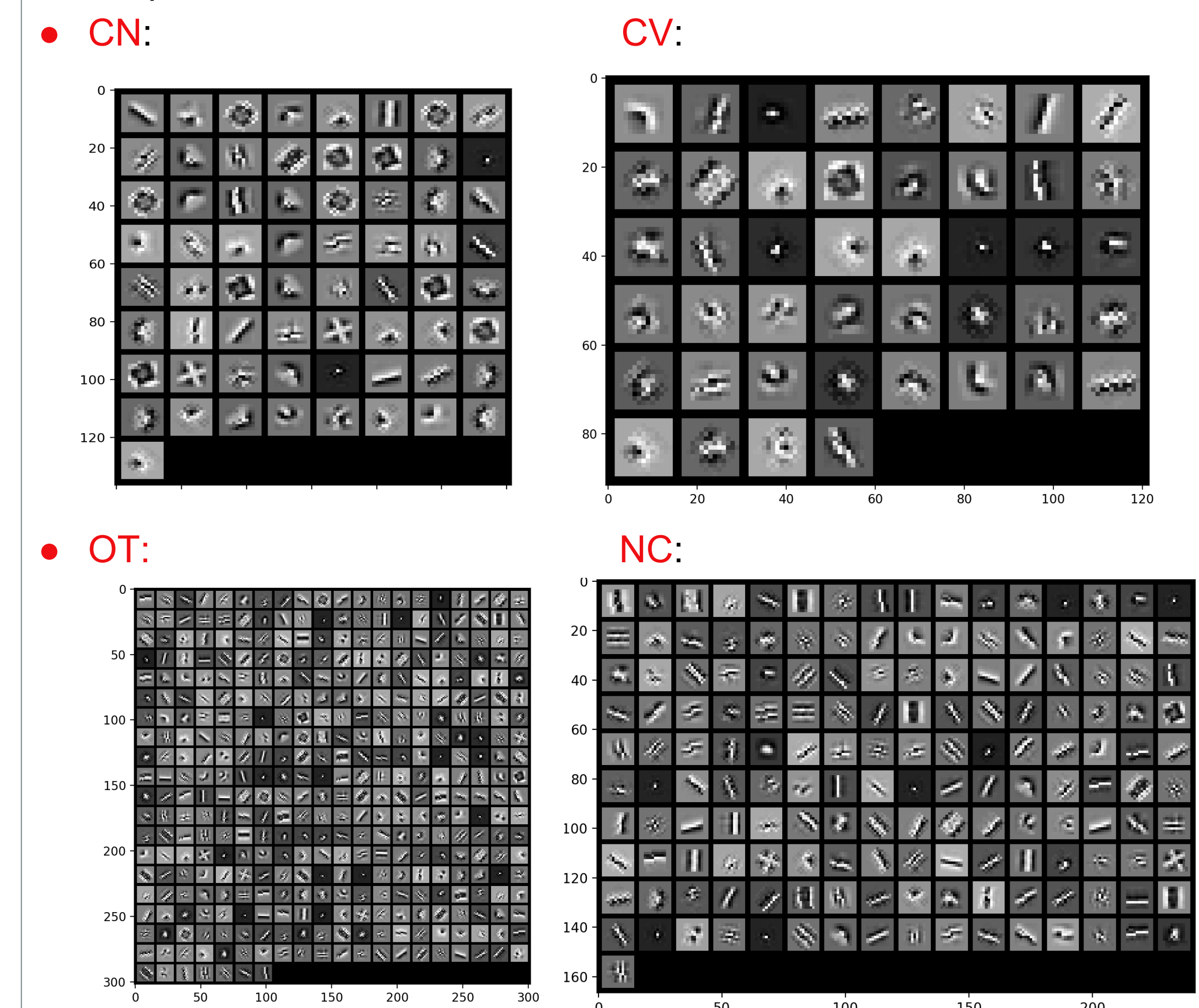


Figure 9: Previous model's Classification result and these Neuron's top filter selected by CMPR method. This provide more than 65% agreement.

Results

- Training performance: CPPR > SMPR > CMPR
 - Testing performance: CMPR > SMPR > CPPR
- SCNN and CNN have similar performance
SCNN is slightly better than CNN

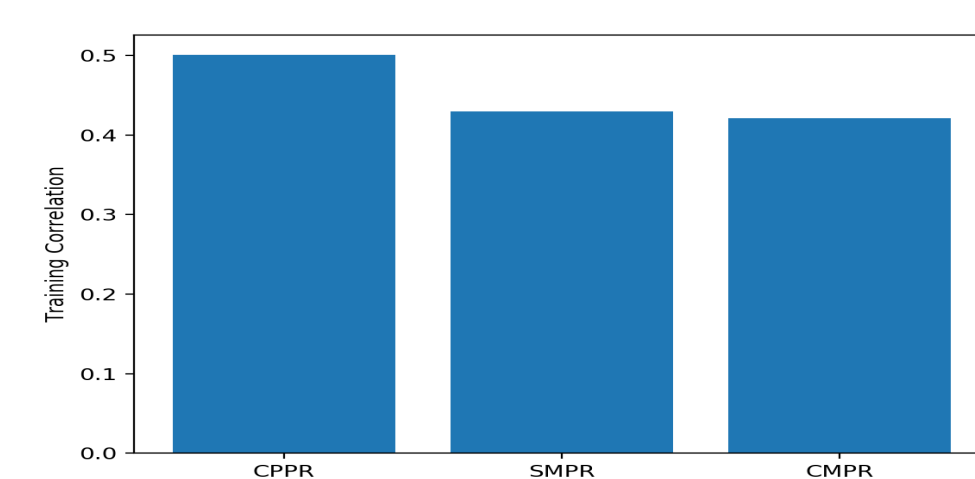


Figure 5: Model Performance of Pursuit models

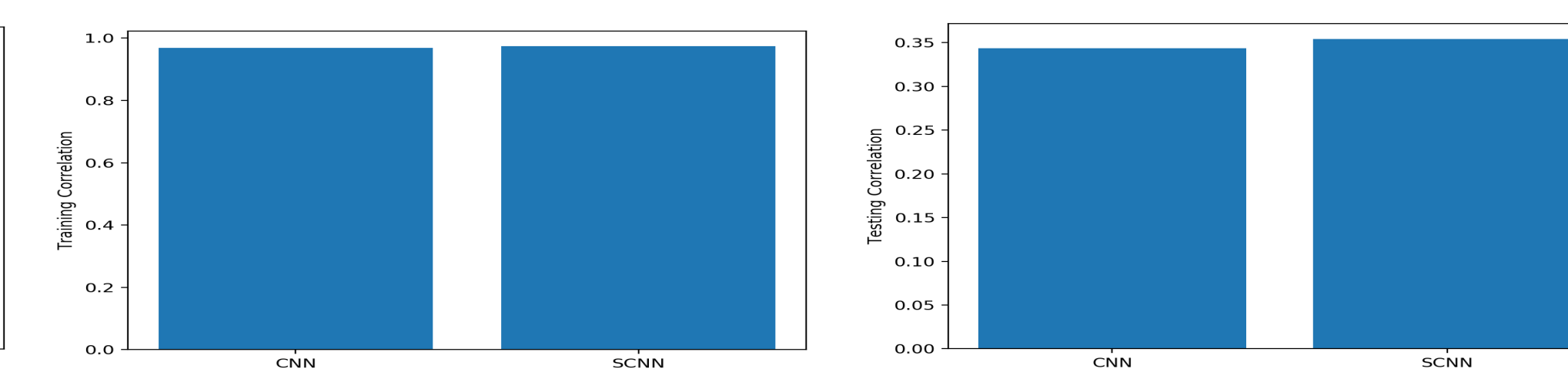


Figure 7: Model Performance of CNN models

- Learnt Filter Example:

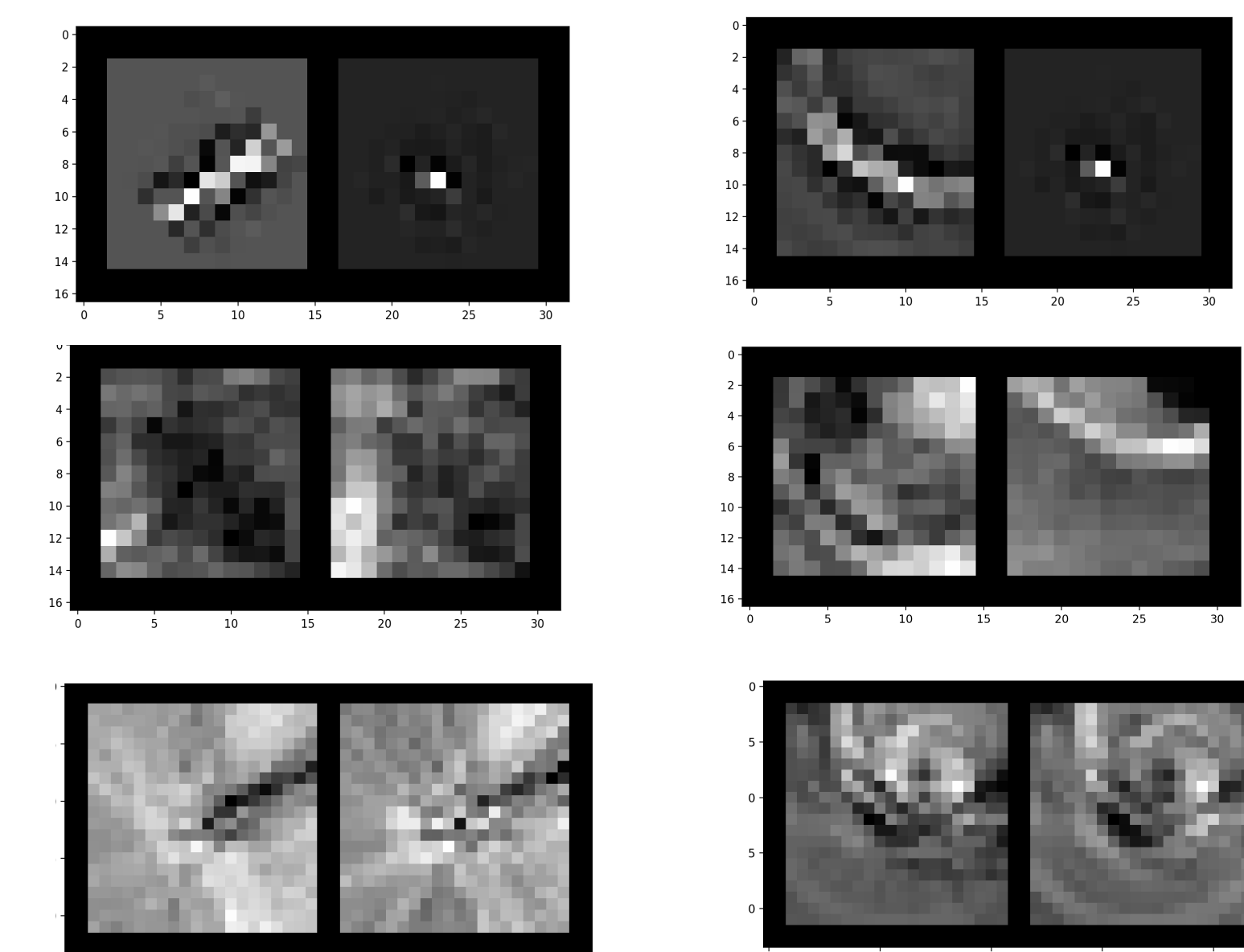


Figure 6: neuron 578 (left) and 553(right). From Top to bottom are the filters of CMPR, CPPR and PPR. Two layer are shown

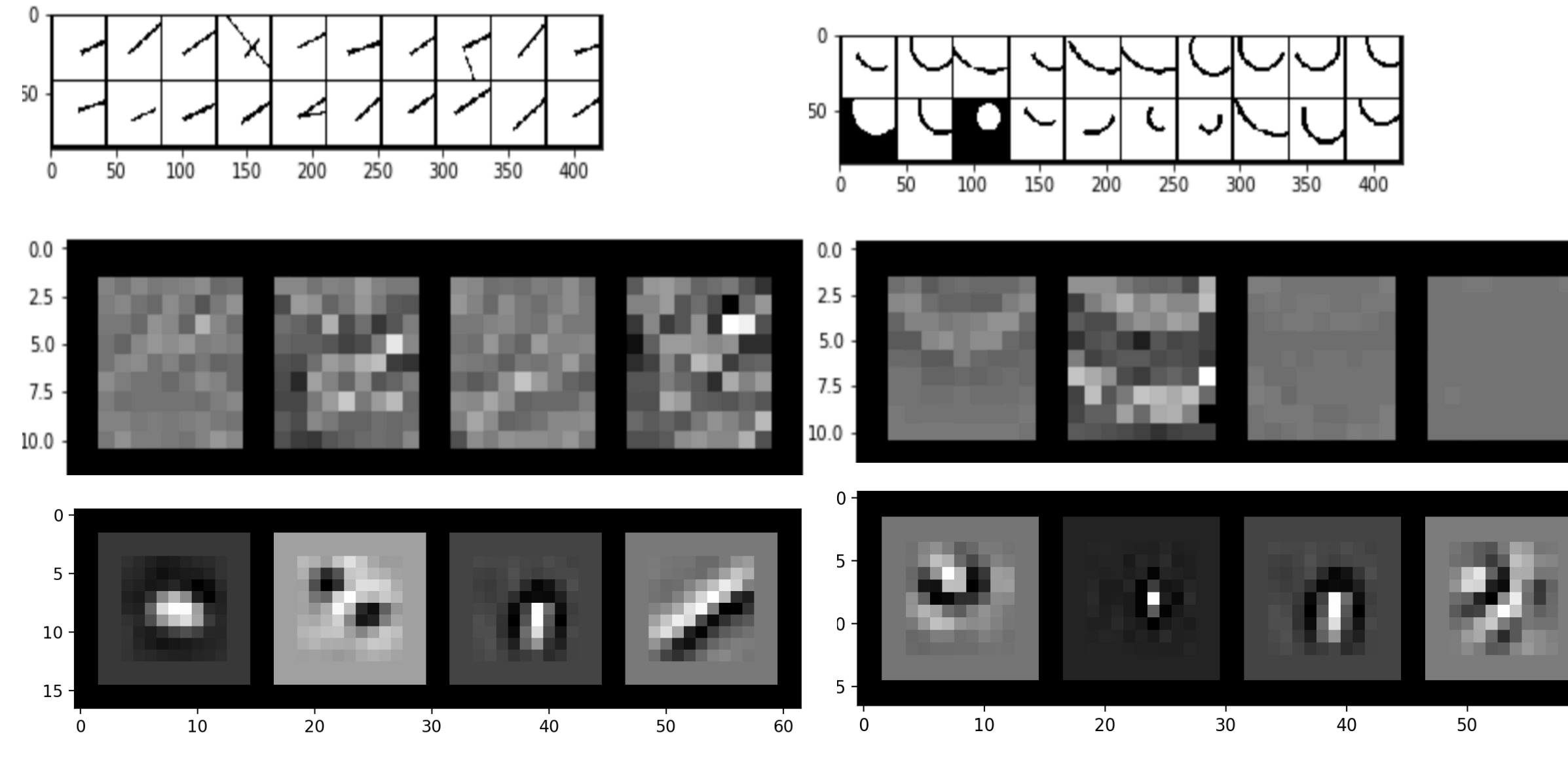


Figure 8: neuron 578 (left) and 553(right). From Top to bottom are the top stimulus, filters of CNN and top weighted filter of SCNN

Conclusion and Future Work

- Only studying the top stimulus of the neuron helps recover clearer neuron's receptive field. On the contrary, traditional method tend to fit too much "noise".
- Incorporating the idea of transfer learning, our model uses the learnt complete set of filters from sparse coding to predict the neuronal response. **This method not only provides a interpretable result but achieves a better performance by reducing overfitting as well.**
- In addition, this method presents a systematic way of classifying visual neurons and roughly agrees with our old method of neuron classification.
- Future Work: Train my model on different filter dictionaries and compare the result. Less complete dictionary with only Gabor-based feature. Over complete filter dictionary.
- Acknowledgement: Dataset is provided by Prof. Tang from Peking University, China.