

Figure 4: Performances of DC-RecNN with the different sizes of latent vector z on five development datasets: IE, MR, SST, SUBJ, and QC. Y-axis represents the accuracy(%), and X-axis represents dimensionality of z .

	Type	Neurons	Examples	Explanations
Semantic	Lexical	17- <i>th</i>	fun, glad, terrific, wonderful, refreshing	Words related to sentiment
	Phrasal	21- <i>st</i>	pick holes, see red, in stitches, split hairs	Phrases related to sentiment
Syntactic	Noun Phrase	45- <i>th</i>	blond boy, pink shirt, green grass, black dog	Containing modifiers related to color
	Verb Phrase	27- <i>th</i>	waking up, take off, pulling up, driving down	Phrases constructed by light-verb
	Prep. Phrase	11- <i>th</i>	slicing a potato, playing guitar, chopping butter	Verb-object phrases
		13- <i>th</i>	on a track, in rocky area, on a stage, over water	Phrases related to places

Table 5: Multiple interpretable neurons and the words/phrases captured by these neurons. The last column gives the explanations of corresponding neuron’s behaviours.

process of composition. Figure 5-(a) shows a visualization. We can see in this sentence, the neuron has realized that this idiomatic collocation “in stitches” is a key pattern, which is crucial for the final sentiment prediction.

- For more complicated tasks such as semantic matching, a well-grounded understanding of the syntactic structure is crucial. In this context, we find that a meta network could capture some syntactic information. For example, the 27-*th* neuron monitors phrases constructed by light-verb. As shown in Figure 5-(b), the verb phrase “taking off” has been attended for forthcoming compositional operation, which is more useful for judging the semantic relation between the sentence pair “An airplane is taking off/A plane is landing”.

6 Related Work

One thread of related work is the exploration of different kinds of compositional function over tree structures. Socher *et al.* [2012] proposed the recursive neural network with standard compositional function. After that, some extensions are introduced to enhance the expressive power of compositional function, such as MV-RecNN [Socher *et al.*, 2013b], SU-RNN [Socher *et al.*, 2013a], RNTN [Socher *et al.*, 2013b], while these models suffer from the problem of hard-coded compositional operations and overfitting.

Another thread of work is the idea of using one network to direct the learning of another network [De Brabandere *et al.*, 2016]. Naik and Mammone [1992] introduce a meta neural network to provide another network with a step size and a direction vector, which is helpful for parameter optimization. De Brabandere *et al.* [2016] propose the dynamic filter

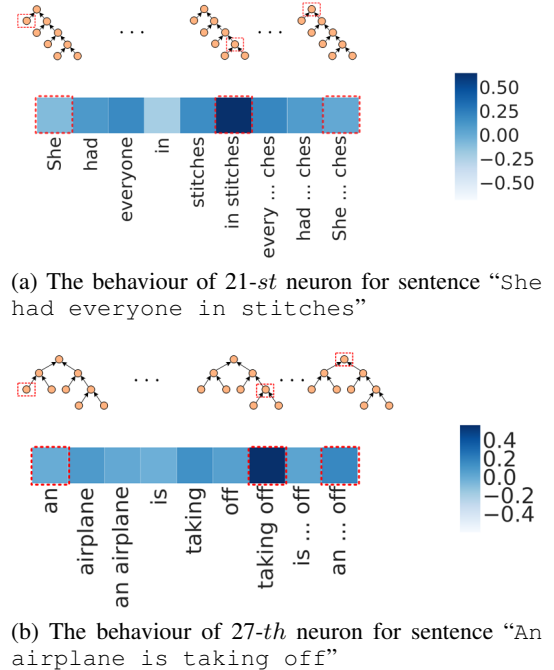


Figure 5: The two heat maps describe the behaviours of neurons z_{21} and z_{27} from DC-TreeNN.

network to implicitly learn a variety of filtering operations. Bertinetto *et al.* [2016] introduce a learnnet for one-shot learning, which can predict the parameters of a second network given a single exemplar. Ha *et al.* [2016] propose the model hypernetwork, which uses a small network to generate the weights for a larger network.

Different from these models, we employ the idea of param-