a)  Tone down the claims that this work (or artificial neural networks in general) achieves performance that is equivalent to human cognition or intelligence, although that may be the long-term vision.

I have modified the following over-claimed sentences:

Abstract:

* Line 1, Page 10.

NE has led to the development of biologically-inspired computer architectures ~~which may~~ whose long-term goal is to approach the performance of the human brain in terms of energy efficiency and cognitive capabilities.

Chapter 1:

* bottom of Page 17

NE has led to the development of biologically-inspired computer architectures ~~which may~~ whose long-term goal is to approach the performance of the human brain in terms of energy efficiency and cognitive capabilities.

* middle paragraph, Page 16

STDP as a learning mechanism based on biological observations has been ~~theoretically proved~~ implemented to be equivalent to a stochastic version of powerful machine learning algorithms.

Chapter 3:

* Section 3.1, Page 45

Deep Learning ~~seems to have~~ has become the answer to ~~all~~ increasing number of artificial intelligence problems ~~overnight since Geoffrey Hinton~~ since Hinton et al. [2006] firstly proposed the training method of ~~a type of ANN,~~ the Deep Belief Network, ~~in 2006 [Hinton et al., 2006]~~.

b)  Include a section on the problems to be solved for spiking networks to be able to learn continuously in practice, e.g. controlling learning, protecting memories (e.g. Grossberg’s stability-plasticity dilemma), segmentation.

At the beginning of Chapter 5 (on-line SNN training), I added a paragraph to state the ‘on-line’ and ‘off-line’ learning in the field of Neuromorphic Engineering. Thus, readers from various background could have a clearer and unified definition of ‘on-line’ system in the context of this thesis.

Then, I included a new paragraph of the stability-plasticity dilemma in artificial systems after stating the brain is a natural on-line system:

One of the practical problems, the stability-plasticity dilemma [Grossberg, 1987] , is a typical example which only exists in artificial systems but not in the brain. Off-line trained systems cannot learn anything new, whereas on-line learning systems easily lose their previous knowledge. However, the brain intuitively achieves both stability and plasticity simultaneously; it maintains gained knowledge while being plastic in respond to new input. Hence, there will be important lessons, such as controlling learning, protecting memories/memory segmentation, and etc., to learn from the brain before an on-line SNN system delivers genuine learning capability.

c)  There are a large number of acronyms in the thesis. Although they are defined on first use, a table including them all is necessary.

I add a list of acronyms before Chapter 1.

d)  Use consistent symbols in the equations throughout the thesis, which should also be defined in another table.

I add a list of mathematical symbols before Chapter 1.

e)  P63 fig 4.2 and other similar figures should be replotted to make the distinction between the various lines clearer. A clearer explanation of what these plots show is also needed

I have modified Figures 4.2, 4.3, 4.5, 4.7, 4.8, 4.9. 4.13, 4.15 with bolder lines in colour and increased image sizes to show a clearer distinction between the curves. I also added text to better explain the figures in the context, page 65-73, and modified the captions of the figures.

f)  A clearer explanation of how and where the offset term which effectively replaced the learnable biases used artificial networks is needed

First of all, I explained why the biases of neurons in artificial neural network is not necessary for the task of MNIST in Chapter 2, page 32.

Usually, a bias is included in the weighted summation which ~~can be seen as an extra input xb = 1 with its weight set to b~~ increases the expression ability of a neuron. However, in this thesis we ~~exclude biases for both artificial and spiking neurons~~ remove biases of both ANNs and SNNs to simplify the neural models and to reduce the number of parameters. Nevertheless, our experimental results show that the performance almost keeps the same when solving a relatively simple task, the MNIST.

Secondly, instead of hard-coding the current bias to 0.1 (i\_offset=0.1 nA for all the LIF neurons, see Table 4.1), I include the current bias as a third parameter of NSP to fit practical LIF firing rates. By setting i\_offset to 0, we can clearly see the curves of LIF simulation is a shifted and scaled NSP in Figure 4.7. Consequently, there are a few places to modify in the Chapter:

1. All the LIF simulations are retested using i\_offset=0, so the Figures 4.2, 4.3, 4.5, 4.7, 4.9 are all shifted on the x-axis.
2. Section 4.4.1 is rewritten to curve-fit the parameters of (k, b, S) of NSP to practical LIF response firing rates.
3. The result section 4.5.2 and Figure 4.12 show better fitting of NSP to LIF response activities, since b is also included for parameter fitting.

g)  Learning in the spiking network needs to be more clearly explained – it is not clear exactly which parameter is learnt in the final implementation

In Chapter 4 of the off-line training method, I added one paragraph in Page 81:

As stated in Section 3.2, the Backpropagation algorithm updates weights using the optimisation method, stochastic gradient descent, to minimise error between the labels and the predictions from the network.

In Chapter 5, I also added one sentence in Page 100:

In another word, the final on-line learning implementation is to accurately set the learning rates ηs+ and ηs− and the time window τdur for the STDP learning rule, and the weights of two connected spiking neuron will be continually modified given synchronous spikes of the pair of neurons.

h)  The transformations for achieving the PAF need to be explained more clearly (p72)

I have modified the Section 4.4.2 mainly from Page 78-80, to better explain the figures, annotate the parameters and claim the transformation process more clearly.

i)  The transformations between the activation function and firing rates need to be explained more carefully so that it is clear how the firing rates are kept within reasonable bounds.

I added one paragraph in Page 80 to explain the transformation from numerical values to firing rates.

j)  P93 and related figures: precisely what the goal of this series of simulations is, what the data is, how to   
read the plots, etc should be explained. The use of the term ‘images’ is misleading.

The goal of the simulations is stated in the first paragraph of the Section 5.4.

I also aded one sentence to make it more clear:

This section attempts to verify the SRM method in practice, thus we compare the learning performance of the conventional Deep Learning models to their spiking versions. We record the reconstruction performance of the models, and carefully observe the dynamics of weights modifications, the activities of the hidden and output neurons and the reconstruction loss.

The data is also clearly described in Section 5.4.1, including how to read the data inputs of Figure 5.5.

The term of ‘image’ is better explained:

thus providing an accurate comparison of the weight updates. The input ~~vector~~ data vector of ten dimensions, seen as an image, repeatedly fed into the network 5,000 times during training. Representing a data vector with the term ‘image’ helps us to better demonstrate the reconstruction task, and to have a unified expression as images in the MNIST dataset for later use.

k)  In chapter 6 make it clear what your contributions were as this was a team effort.

In Section 6.3 and the conclusion, I declare the work done other people and the contribution I made.

l)  Please see both copies of the thesis which the examiners have annotated for further minor points of corrections

I have finished all the corrections annotated on the copies of the thesis.