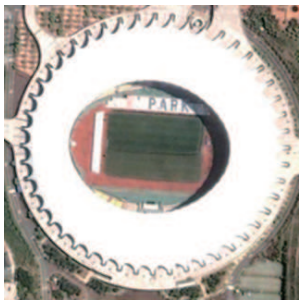
A Corpus Sample

Remote sensing images refer to the photography which record the size of electromagnetic wave of various ground objects, and are usually taken by satellite. Samples are shown in Figure 3. For (a), the house is a typical ground object, and while we ask "What is on the field?", we get "A house" as answer.

(a) Q: What is on the field?
A: A house.



(b) Q: What is the shape of the stadium?
A: Circle.

(c) Q: How many cars are there on the bridge?
A: Two.

(d) Q: Are there any playgrounds?
A: Yes.

Figure 3: Four remote sensing images attached with one QA pair sampled from our remote sensing question answering corpus: (a) sparse residual area; (b) stadium; (c) bridge; (d) resort.

B Distribution of Questions by Their First Four Words

We illustrate distribution of questions by their first four words in Figure 4.

C Top Frequent Answers

We show top frequent answers from our remote sensing question answering corpus in Figure 5.

D Line Chart on the Variation Trend in Terms of Accuracy

We show the variation trend in terms of accuracy with epoch growth in Figure 6. From the line chart we observe that the accuracy changes within the range of 29.2% to 30.2%. Moreover, the change tends to be gentle.

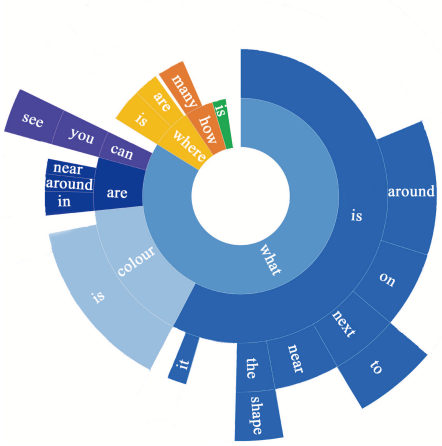


Figure 4: Distribution of questions by their first four words. The ordering of the words starts towards the center and radiates outwards. The angles of the regions are proportional to the number of questions containing the word. White areas are words with contributions too small to show.

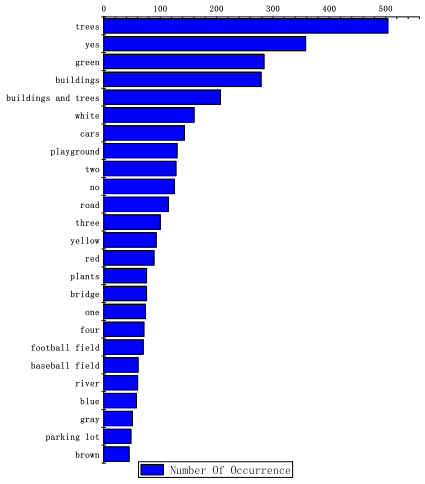


Figure 5: Histogram of top frequent answers from our remote sensing question answering corpus.

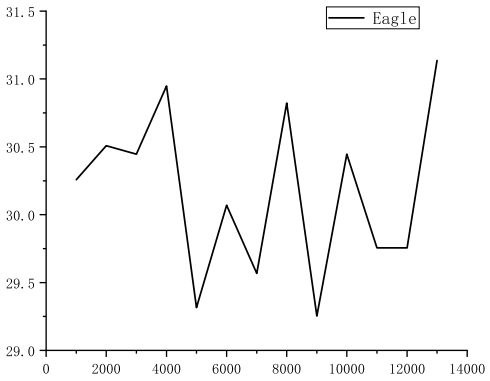


Figure 6: Line Chart on the Variation Trend in Terms of Accuracy.