|  |  |  |
| --- | --- | --- |
|  | Xi | Yi |
| Image feature vector | Partial caption | Target word |
| img1 | startseq | i |
| img1 | startseq i | love |
| img1 | startseq I love | studying |
| img1 | startseq I love studying | natural |
| img1 | startseq I love studying natural | language |
| img1 | startseq I love studying natural language | processing |
| img1 | startseq I love studying natural language processing | endseq |

|  |  |  |
| --- | --- | --- |
|  | Xi | Yi |
| Image feature vector | Partial caption | Target word |
| img1 | [7, 0, 0 ..., ..., ..., ..., ..., ..., 0] | 2 |
| img1 | [7, 2, 0, 0 ..., ..., ..., ..., ..., 0] | 4 |
| img1 | [7, 2, 4, 0, 0 ..., ..., ..., ..., 0] | 8 |
| img1 | [7, 2, 4, 8, 0, 0 ..., ..., ..., 0] | 5 |
| img1 | [7, 2, 4, 8, 5, 0, 0 ..., ..., 0] | 3 |
| img1 | [7, 2, 4, 8, 5, 3, 0, 0 ..., ..., 0] | 6 |
| img1 | [7, 2, 4, 8, 5, 3, 6, 0, 0 ..., 0] | 1 |