|  |  |  |
| --- | --- | --- |
|  | Number of Learnable Parameters | Train accuracy |
| Width -1 and depth -1 | 1,402,932 | 84 % |
| Width -1 and depth -2 | 2,973,492 | 85% |
| Width -1 and depth -3 | 4,544,052 | 86% |
| Width -2 and depth -1 | 5,595,220 | 89% |
| Width -3 and depth -1 | 12,576,884 | 90% |