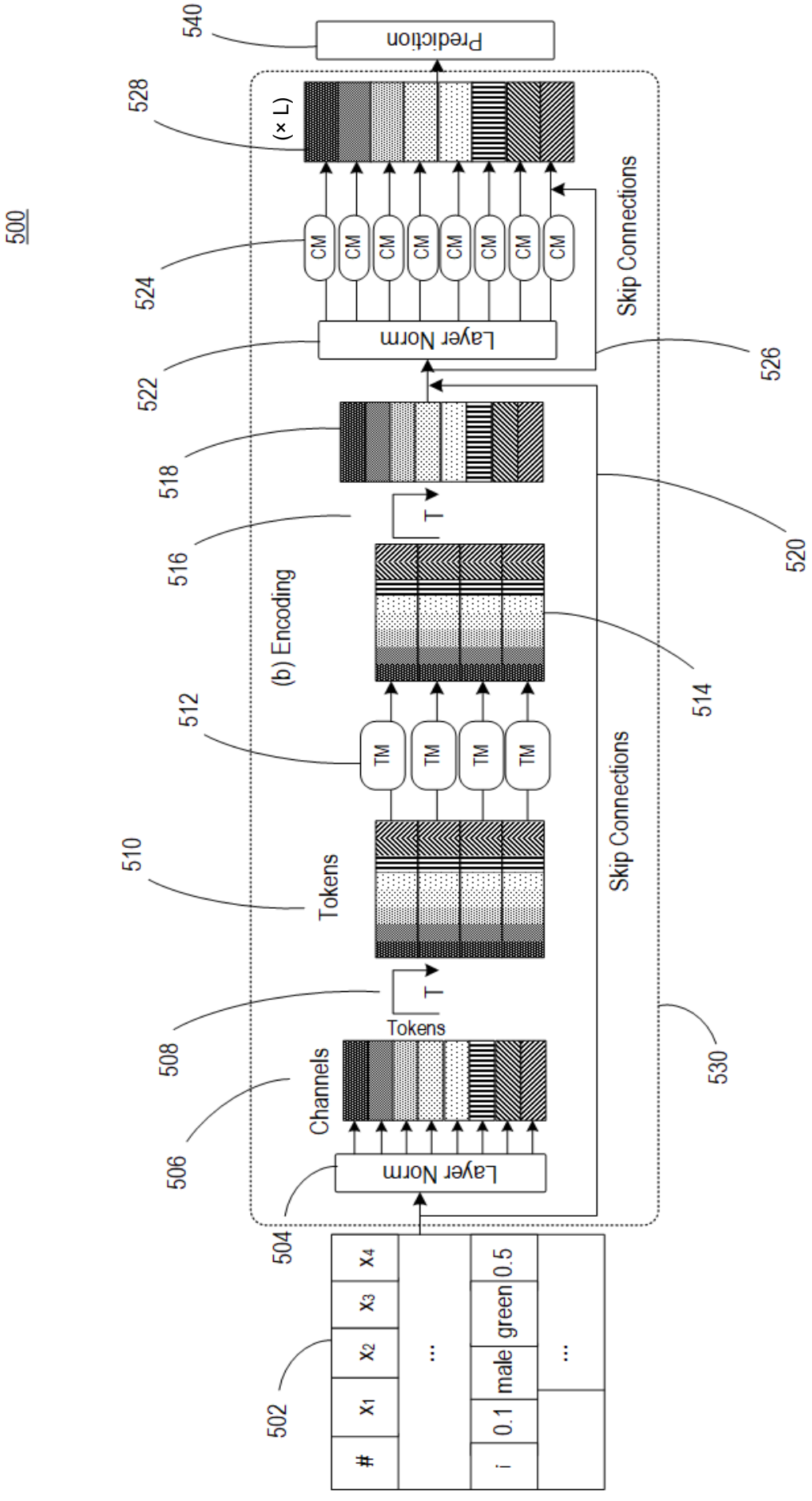

500
502
504
506
508
510
512
514
516
518
520
522
524
526
528
530
540
#
x1
x2
x3
x4
...
i
0.1
male
green
0.5
...
Layer Norm
Channels
Tokens
T
Tokens
TM
TM
TM
TM
(b) Encoding
T
Skip Connections
Skip Connections
Layer Norm
CM
CM
CM
CM
CM
CM
CM
CM
(× L)
Prediction

FIG. 5