Model performance in CER (%) against each dialect with pre-training 80 71.51 66.34 47.79 63.84 78.79 Egyptian - $(\pm 6.01)$  $(\pm 3.37)$  $(\pm 4.04)$  $(\pm 10.96)$  $(\pm 7.65)$ 75 53.18 61.76 64.08 54.64 68.79 Gulf - $(\pm 1.37)$  $(\pm 4.13)$  $(\pm 2.89)$  $(\pm 0.5)$  $(\pm 1.77)$ - 70 Dialect fine-tuned on 61.08 67.3 66.02 81.54 54.74 65 Iraqi - $(\pm 4.37)$  $(\pm 3.59)$  $(\pm 6.33)$  $(\pm 8.82)$  $(\pm 13.14)$ 60 51.03 60.61 60.33 49.1 65.64 Levantine - $(\pm 3.3)$  $(\pm 6.31)$  $(\pm 4.96)$  $(\pm 4.25)$  $(\pm 6.12)$ 55 55.48 65.24 63.73 58.15 61.67 Maghrebi - $(\pm 5.54)$  $(\pm 7.93)$  $(\pm 5.7)$  $(\pm 9.75)$  $(\pm 12.95)$ - 50 Dialect tested on