Average validation loss during training on strict early stopping parameters

- Run with batch size = 128, weight decay = 5e-05, momentum = 0.97
- Run with batch size = 64, weight decay = 5e-05, momentum = 0.96
- Run with batch size = 32, weight decay = 5e-05, momentum = 0.96

(29, 0.0394)

(30, 0.0201)

(37, 0.0103)

Validation loss

Epochs