Generalization with SGD Training