**DEEP LEARNING IN COMPUTER VISION - EXERCISE 1**

KARSTEN HERTH, FELIX HUMMEL, FELIX KAMMERLANDER, AND DAVID PALOSCH

**Exercise 1.2 (e)**
The model is not able to remove the noise. This is due to the fact, that the model tries to reduce vignetting according to a radial symmetric function. However, the noise does not behave according to a radial symmetric law. In practice, we think that noise even gets worse after applying devignetting to the images. This could be explained as follows: If the noise was added to the pictures after vignetting with some standard derivation $\sigma > 0$ and we apply the model in order to remove the vignette, then the noise gets scaled by the factor we applied, i.e. the observed standard derivation $\sigma'$ after removing the vignetting has increased compared to the original noise, i.e. $\sigma' > \sigma$.

**Exercise 1.3**

(a) Let $E\colon \mathbb{R}^3 \to \mathbb{R}$ be given by

$$E(\theta) := 2\theta_1^2 + 4\theta_2 + \max\{0, \theta_2 + \theta_3\}.$$

Performing two steps of gradient descent for $E$ starting at $\theta^{(0)} = (2,1,0)^\top$ with step size $\tau = \frac{1}{2}$, we end up at $\theta^{(2)} = (x,x,x)^\top$.

(b) Let $f\colon \mathbb{R}^n \to \mathbb{R}$ and $g\colon \mathbb{R} \to \mathbb{R}$ both be differentiable and $h := g \circ f$. Then,

$$\nabla h(p) = g'(f(p))\nabla f(p) \quad (p \in \mathbb{R}^n).$$

*Proof.*

(a) The gradient of $E$ is given by

$$\nabla E(\theta) = \begin{cases} (4\theta_1, 5, 1)^\top, & \theta_2 + \theta_3 \geq 0, \\ (4\theta_1, 4, 0)^\top, & \text{else.} \end{cases}$$

Hence, we have $\nabla E(\theta^{(0)}) = (8,5,1)^\top$ and we obtain

$$\begin{aligned} \theta^{(1)} &= \theta^{(0)} - \tau \nabla E(\theta^{(0)}) \\ &= (2,1,0)^\top - \frac{1}{2}(8,5,1)^\top \\ &= (-2, -\tfrac{3}{2}, -\tfrac{1}{2})\top. \end{aligned}$$

This time, holds that $\nabla E(\theta^{(1)}) = (-8,4,0)^\top$ and consequently,

$$\begin{aligned} \theta^{(2)} &= (-2, -\tfrac{3}{2}, -\tfrac{1}{2})^\top - \frac{1}{2}(-8,4,0)^\top \\ &= (2, -\tfrac{7}{2}, -\tfrac{1}{2})^\top. \end{aligned}$$

(b) Let $q \in \mathbb{R}^n \setminus \{0\}$. Then

$$\begin{aligned} h(p+q) - h(p) &= g(f(p+q)) - g(f(p)) \\ &= g(f(p) + \nabla f(p)q + o(|q|)) - g(f(p)) \\ &= g(f(p) + z_q) \end{aligned}$$

*Date*: November 12, 2017.

holds for all $p \in \mathbb{R}^n$ where we have set $z_q := \nabla f(p)q + o(|q|) \in \mathbb{R}$. Using the differntiability of $g$ at $f(p)$ we obtain

$$\begin{aligned} h(p+q) - h(p) &= g(f(p)) + g'(f(p))z_q + o(|z_q|) \\ &= h(p) + g'(f(p))\nabla f(p)q + g'(f(p))o(|q|) + o(|z_q|) \\ &= h(p) + g'(f(p))\nabla f(p)q + o(|q|) \end{aligned}$$

for all $p \in \mathbb{R}^n$, since $o(|z_q|) = o(|q|)$. Hence, the assertion follows.

$\square$