Supplementary information for "*Beyond MixUps: Task-Guided Variational Partial Synthesis for Fundus Image Analysis*"

Enyi Li, Gaoruishu Long, Xuhuan Wang, Wei Chen, and Jinchao Liu, *Member, IEEE*

I. DERIVATION OF THE ELBO LOSS FOR TVPS

Here we show how to derive the *evidence lower bound (ELBO)* of our proposed model in details. We started from the marginal likelihood of observable variables (x, y)

$$
\begin{aligned}
\log p(\boldsymbol{x}, \boldsymbol{y}) &= \mathbb{E}_{p(\boldsymbol{z}, \boldsymbol{\omega}|\boldsymbol{x}, \boldsymbol{y})}[\log p(\boldsymbol{x}, \boldsymbol{y})] \\
&= \mathbb{E}_{p(\boldsymbol{z}, \boldsymbol{\omega}|\boldsymbol{x}, \boldsymbol{y})}\left[\log\left[\frac{p(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\omega})}{p(\boldsymbol{z}, \boldsymbol{\omega}|\boldsymbol{x}, \boldsymbol{y})}\right]\right] + \mathrm{KL}(p(\boldsymbol{z}, \boldsymbol{\omega}|\boldsymbol{x}, \boldsymbol{y})||p(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\omega}))
\end{aligned}
$$

where $\boldsymbol{x}, \boldsymbol{y}$ are data samples of input and output. $\boldsymbol{z}$ denotes the latent variables, the posterior of which is defined to be mixture of Gaussian. The mixture of Gaussian has K Gaussian components and the mixing coefficients $\boldsymbol{s} = \{s_1, \cdots, s_K\}$. We denote the corresponding latent random variable to determine which Gaussian component to draw sample from as ω.

As discussed in the main text, we parameterize the posterior $p(\boldsymbol{z}, \boldsymbol{\omega}|\boldsymbol{x})$ of latent variables $\boldsymbol{z}, \omega$ using a neural encoder F_ϕ. We denote the posterior as $q_\phi(\boldsymbol{z}, \boldsymbol{\omega}|\boldsymbol{x})$. To decode representations sampled from $\boldsymbol{z}$, we design a decoder G_θ which takes a random sample and output a reconstruction $\boldsymbol{y}$, with additional information from the encoder. Therefore, unlike regular VAE, our generative model is a composition of both F_ϕ and G_θ. We train F_ϕ and G_θ by maximizing the marginal likelihood of observable variables $(\boldsymbol{x}, \boldsymbol{y})$.

$$
\begin{aligned}
\log p(\boldsymbol{x}, \boldsymbol{y}) &= \mathbb{E}_{q_\phi(\boldsymbol{z}, \boldsymbol{\omega}|\boldsymbol{x})}[\log p_{\theta, \phi}(\boldsymbol{x}, \boldsymbol{y})] \\
&= \mathbb{E}_{q_\phi(\boldsymbol{z}, \boldsymbol{\omega}|\boldsymbol{x})}\left[\log\left[\frac{p_{\theta, \phi}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\omega})}{q_\phi(\boldsymbol{z}, \boldsymbol{\omega}|\boldsymbol{x})}\right]\right] + \mathrm{KL}(q_\phi(\boldsymbol{z}, \boldsymbol{\omega}|\boldsymbol{x})||p_{\theta, \phi}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\omega}))
\end{aligned}
$$

The ELBO thus reads:

$$
\begin{aligned}
\mathcal{L}_{\theta, \phi}(\boldsymbol{x}, \boldsymbol{y}) &= \mathbb{E}_{q_\phi(\boldsymbol{z}, \boldsymbol{\omega}|\boldsymbol{x})}\left[\log\left[\frac{p_{\theta, \phi}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\omega})}{q_\phi(\boldsymbol{z}, \boldsymbol{\omega}|\boldsymbol{x})}\right]\right] \\
&= \mathbb{E}_{q_\phi(\boldsymbol{z}, \boldsymbol{\omega}|\boldsymbol{x})}\left[\log\left[\frac{p_{\theta, \phi}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\omega})}{q_\phi(\boldsymbol{z}, \boldsymbol{\omega}|\boldsymbol{x})}\right]\right] \\
&= \mathbb{E}_{q_\phi(\boldsymbol{z}, \boldsymbol{\omega}|\boldsymbol{x})}\left[\log p_{\theta, \phi}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\omega}) - \log q_\phi(\boldsymbol{z}, \boldsymbol{\omega}|\boldsymbol{x})\right] \\
&= \mathbb{E}_{q_\phi(\boldsymbol{z}, \boldsymbol{\omega}|\boldsymbol{x})}\left[\log p_{\theta, \phi}(\boldsymbol{x}, \boldsymbol{y}|\boldsymbol{z}) + \log p(\boldsymbol{z}|\boldsymbol{\omega}) + \log p(\boldsymbol{\omega}) - \log q_\phi(\boldsymbol{z}, \boldsymbol{\omega}|\boldsymbol{x})\right] \\
&= \mathbb{E}_{q_\phi(\boldsymbol{z}, \boldsymbol{\omega}|\boldsymbol{x})}\left[\log p_{\theta, \phi}(\boldsymbol{x}, \boldsymbol{y}|\boldsymbol{z})\right] + \mathbb{E}_{q_\phi(\boldsymbol{z}, \boldsymbol{\omega}|\boldsymbol{x})}\left[\log p(\boldsymbol{z}, \boldsymbol{\omega}) - \log q_\phi(\boldsymbol{z}, \boldsymbol{\omega}|\boldsymbol{x})\right] \\
&= \underbrace{\mathbb{E}_{q_\phi(\boldsymbol{z}, \boldsymbol{\omega}|\boldsymbol{x})}\left[\log p_{\theta, \phi}(\boldsymbol{x}, \boldsymbol{y}|\boldsymbol{z})\right]}_{\text{reconstruction loss}} + \underbrace{\mathrm{KL}(q_\phi(\boldsymbol{z}, \boldsymbol{\omega}|\boldsymbol{x})||p(\boldsymbol{z}, \boldsymbol{\omega}))}_{\text{regularization of latent distributions}} \\
&= \mathbb{E}_{q_\phi(\boldsymbol{z}, \boldsymbol{\omega}|\boldsymbol{x})}\left[\log p_{\theta, \phi}(\boldsymbol{x}, \boldsymbol{y}|\boldsymbol{z})\right] + \mathbb{E}_{q_\phi(\boldsymbol{z}, \boldsymbol{\omega}|\boldsymbol{x})}\left[\log p(\boldsymbol{z}, \boldsymbol{\omega})\right] - \mathbb{E}_{q_\phi(\boldsymbol{z}, \boldsymbol{\omega}|\boldsymbol{x})}\left[\log q_\phi(\boldsymbol{z}, \boldsymbol{\omega}|\boldsymbol{x})\right] \\
&= \mathcal{L}_{logxy|z\omega} + \mathcal{L}_{logz\omega} - \mathcal{L}_{logz\omega|x}
\end{aligned}
$$

The ELBO consists of three terms, where the first one is a log-likelihood loss which measures the reconstruction error, the second and third ones correspond to the priors and posterior of the latent variables $\boldsymbol{z}$ and $\boldsymbol{\omega}$.

$$
\mathcal{L}_{logxy|z\omega} = \mathbb{E}_{\boldsymbol{z}, \boldsymbol{\omega}}\left[\log p_{\theta, \phi}(\boldsymbol{x}, \boldsymbol{y}|\boldsymbol{z})\right] \approx \frac{1}{N}\sum_{n=1}^{N}\log p_{\theta, \phi}(\boldsymbol{x}, \boldsymbol{y}|\mathring{\boldsymbol{z}}_{\boldsymbol{n}}, \mathring{\boldsymbol{\omega}}_{\boldsymbol{n}})
$$

where $\{\mathring{z}_n, \mathring{\omega}_n, n = 1, \cdots, N\}$ are random samples drawn from $\mathcal{N}(0, \boldsymbol{I})$ and $\mathcal{M}(K)$.

We assume the prior of z is a standard Gaussian distribution $\mathcal{N}(0, \boldsymbol{I})$.

$$
\begin{aligned}
\mathcal{L}_{logz\omega} &= \mathbb{E}_{\boldsymbol{z},\omega \sim q_\phi(\boldsymbol{z},\boldsymbol{\omega}|\boldsymbol{x})}\Big[\log p(\boldsymbol{z},\boldsymbol{\omega})\Big] \\
&= \mathbb{E}_{\boldsymbol{\omega} \sim q_\phi(\boldsymbol{z},\boldsymbol{\omega}|\boldsymbol{x})}\bigg[\int_{\boldsymbol{z}} q_\phi(\boldsymbol{z},\boldsymbol{\omega}|\boldsymbol{x})\log p(\boldsymbol{z},\boldsymbol{\omega})d\boldsymbol{z}\bigg] \\
&= \mathbb{E}_{\boldsymbol{\omega} \sim q_\phi(\boldsymbol{z},\boldsymbol{\omega}|\boldsymbol{x})}\bigg[\int_{\boldsymbol{z}} \frac{1}{(2\pi)^{D/2}}\frac{1}{\prod_i^D \sigma_{ki}}\boldsymbol{s}_k^{\omega_k}\exp\Big\{-\frac{1}{2}\sum_i \frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2}\Big\}\log\Big(\frac{1}{(2\pi)^{D/2}}\boldsymbol{s}_0^{\omega_k}\exp\Big\{-\frac{1}{2}\sum_i z_{ki}^2\Big\}\Big)d\boldsymbol{z}\bigg] \\
&= \mathbb{E}_{\boldsymbol{\omega} \sim q_\phi(\boldsymbol{z},\boldsymbol{\omega}|\boldsymbol{x})}\bigg[\int_{\boldsymbol{z}} \frac{1}{(2\pi)^{D/2}}\frac{1}{\prod_i^D \sigma_{ki}}\boldsymbol{s}_k^{\omega_k}\exp\Big\{-\frac{1}{2}\sum_i \frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2}\Big\}\Big[\log\Big(\frac{1}{(2\pi)^{D/2}}\boldsymbol{s}_0^{\omega_k}\Big)-\frac{1}{2}\sum_i z_{ki}^2\Big]d\boldsymbol{z}\bigg] \\
&= \mathbb{E}_{\boldsymbol{\omega} \sim q_\phi(\boldsymbol{z},\boldsymbol{\omega}|\boldsymbol{x})}\bigg[\underbrace{\frac{1}{(2\pi)^{D/2}}\frac{1}{\prod_i^D \sigma_{ki}}\boldsymbol{s}_k^{\omega_k}\log\Big(\frac{1}{(2\pi)^{D/2}}\boldsymbol{s}_0^{\omega_k}\Big)\int_{\boldsymbol{z}}\exp\Big\{-\frac{1}{2}\sum_i \frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2}\Big\}d\boldsymbol{z}}_{Part1} \\
&\quad +\underbrace{\frac{1}{(2\pi)^{D/2}}\frac{1}{\prod_i^D \sigma_{ki}}\boldsymbol{s}_k^{\omega_k}\int_{\boldsymbol{z}}\Big(-\frac{1}{2}\sum_i z_{ki}^2\exp\Big\{-\frac{1}{2}\sum_i \frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2}\Big\}\Big)d\boldsymbol{z}}_{Part2}\bigg]
\end{aligned}
$$

To simplify *the integral terms* in Part1 and Part2, we can obtain:

Part 1:

$$
\begin{aligned}
\int_{\boldsymbol{z}}\exp\Big\{-\frac{1}{2}\sum_i \frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2}\Big\}d\boldsymbol{z} &= \Big(\prod_i^D \sigma_{ki}\Big)\int_{\boldsymbol{z}}\exp\Big\{-\frac{1}{2}\sum_i \varepsilon_{ki}^2\Big\}d\varepsilon_{k1}d\varepsilon_{k2}\cdots\varepsilon_{kD} \\
&= \Big(\prod_i^D \sigma_{ki}\Big)(2\pi)^{D/2}
\end{aligned}
$$

Part 2:

We will decompose the D-dimensional summation into the form of the *D-th* term plus a *D-1* dimensional summation, and solve it recursively.

$$
\begin{aligned}
\int_{\boldsymbol{z}}\Big(-\frac{1}{2}\sum_i z_{ki}^2\exp\Big\{-\frac{1}{2}\sum_i \frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2}\Big\}\Big)d\boldsymbol{z} &= \int_{\boldsymbol{z}}-\frac{1}{2}z_{kD}^2\exp\Big\{\Big(-\frac{1}{2}\sum_i \frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2}\Big)\Big\}dz_{kD}dz_{k(D-1)}\cdots dz_{k1} \\
&\quad +\int_{\boldsymbol{z}}-\frac{1}{2}\sum_i^{D-1} z_{ki}^2\exp\Big\{\Big(-\frac{1}{2}\sum_i \frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2}\Big)\Big\}dz_{kD}dz_{k(D-1)}\cdots dz_{k1}
\end{aligned}
$$

It can be derived that:

$$
\int_{\boldsymbol{z}}\sigma_{kD}^2\varepsilon_{kD}^2\exp\Big\{-\frac{1}{2}\varepsilon_{kD}^2\Big\}d\varepsilon_{kD} = (2\pi)^{1/2}\sigma_{kD}^2 \tag{1}
$$

$$
\int_{\boldsymbol{z}}\mu_{kD}^2\exp\Big\{-\frac{1}{2}\varepsilon_{kD}^2\Big\}d\varepsilon_{kD} = (2\pi)^{1/2}\mu_{kD}^2 \tag{2}
$$

$$
\int_{\boldsymbol{z}}2\mu_{kD}\sigma_{kD}\varepsilon_{kD}\exp\Big\{-\frac{1}{2}\varepsilon_{kD}^2\Big\}d\varepsilon_{kD} = -2\mu_{kD}\sigma_{kD}\int_{\boldsymbol{z}}d\exp\Big\{-\frac{1}{2}\varepsilon_{kD}^2\Big\} = 0 \tag{3}
$$

Then the *D-th* term can be expressed as:

$$
\int_{\boldsymbol{z}}-\frac{1}{2}z_{kD}^2\exp\Big\{\Big(-\frac{1}{2}\sum_i \frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2}\Big)\Big\}dz_{kD}dz_{k(D-1)}\cdots dz_{k1} \tag{4}
$$

$$
= \int_{\boldsymbol{z}}\exp\Big\{\Big(-\frac{1}{2}\sum_i^{D-1} \frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2}\Big)\Big\}\Big(-\frac{1}{2}z_{kD}^2\Big)\exp\Big\{\Big(-\frac{1}{2}\frac{(z_{kD}-\mu_{kD})^2}{\sigma_{kD}^2}\Big)\Big\}dz_{kD}dz_{k(D-1)}\cdots dz_{k1} \tag{5}
$$

$$
= \int_{\boldsymbol{z}}\exp\Big\{\Big(-\frac{1}{2}\sum_i^{D-1} \frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2}\Big)\Big\}\Big(-\frac{\sigma_{kD}}{2}\Big)(\sigma_{kD}\varepsilon_{kD}+\mu_{kD})^2\exp\Big\{-\frac{1}{2}\varepsilon_{kD}^2\Big\}d\varepsilon_{kD}dz_{k(D-1)}\cdots dz_{k1} \tag{6}
$$

$$= -\frac{\sigma_{kD}}{2}(2\pi)^{1/2}(\mu_{kD}^2 + \sigma_{kD}^2)\int_{\boldsymbol{z}} \exp\Big\{(-\frac{1}{2}\sum_i^{D-1}\frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2})\Big\}dz_{k(D-1)}\cdots dz_{k1} \tag{7}$$

$$= -\frac{\sigma_{kD}}{2}(2\pi)^{1/2}(\mu_{kD}^2 + \sigma_{kD}^2)\Big(\prod_{i=1}^{D-1}\sigma_{ki}\Big)(2\pi)^{(D-1)/2} \tag{8}$$

$$= -\frac{1}{2}(\sigma_{kD}^2 + \mu_{kD}^2)\Big(\prod_{i=1}^{D}\sigma_{ki}\Big)(2\pi)^{D/2} \tag{9}$$

And the *D-1* dimensional summation can be expressed as:

$$\int_{\boldsymbol{z}} -\frac{1}{2}\sum_i^{D-1} z_{ki}^2 \exp\Big\{(-\frac{1}{2}\sum_i\frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2})\Big\}dz_{kD}dz_{k(D-1)}\cdots dz_{k1} \tag{10}$$

$$= \int_{\boldsymbol{z}} -\frac{1}{2}\sum_i^{D-1} z_{ki}^2 \exp\Big\{(-\frac{1}{2}\sum_i^{D-1}\frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2})\Big\}\exp\Big\{(-\frac{1}{2}\frac{(z_{kD}-\mu_{kD})^2}{\sigma_{kD}^2})\Big\}dz_{kD}dz_{k(D-1)}\cdots dz_{k1} \tag{11}$$

$$= \sigma_{kD}(2\pi)^{1/2}\int_{\boldsymbol{z}} -\frac{1}{2}\sum_i^{D-1} z_{ki}^2 \exp\Big\{(-\frac{1}{2}\sum_i^{D-1}\frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2})\Big\}dz_{k(D-1)}\cdots dz_{k1} \tag{12}$$

Letting $x_D = \int_{\boldsymbol{z}}\Big(-\frac{1}{2}\sum_i z_{ki}^2 \exp\Big\{-\frac{1}{2}\sum_i\frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2}\Big\}\Big)dz_{kD}dz_{k(D-1)}\cdots dz_{k1}$, we have:

$$x_D = -\frac{1}{2}(\sigma_{kD}^2 + \mu_{kD}^2)\Big(\prod_{i=1}^{D}\sigma_{ki}\Big)(2\pi)^{D/2} + \sigma_{kD}(2\pi)^{1/2}x_{D-1} \tag{13}$$

It can be observed that x_D forms a geometric sequence, then substituting $D = 1$ and applying the formula for the sum of a geometric series, we can obtain:

$$\tag{14}$$

$$x_1 = -\frac{1}{2}(\sigma_{k1}^2 + \mu_{k1}^2)\sigma_{k1}(2\pi)^{1/2} \tag{15}$$

$$x_D = -\frac{1}{2}\Big(\prod_{i=1}^{D}\sigma_{ki}\Big)(2\pi)^{D/2}\Big[\sum_i^{D}\sigma_{ki}^2 + \sum_i^{D}\mu_{ki}^2\Big] \tag{16}$$

Therefore, we can obtain:

$$\mathcal{L}_{logz\omega} = \mathbb{E}_{\boldsymbol{\omega}\sim q_\phi(\boldsymbol{z},\boldsymbol{\omega}|\boldsymbol{x})}\Bigg[\Big(\boldsymbol{s}_k^{\omega_k}\log\Big(\frac{1}{(2\pi)^{D/2}}\boldsymbol{s}_0^{\omega_k}\Big) - \frac{1}{2}\boldsymbol{s}_k^{\omega_k}\Big(\sum_i^{D}\sigma_{ki}^2 + \sum_i^{D}\mu_{ki}^2\Big)\Big)\Bigg] \tag{17}$$

$$= \mathbb{E}_{\boldsymbol{\omega}\sim q_\phi(\boldsymbol{z},\boldsymbol{\omega}|\boldsymbol{x})}\Bigg[\boldsymbol{s}_k^{\omega_k}\Big(\log\frac{1}{K} - \frac{D}{2}\log 2\pi - \frac{1}{2}\sum_i^{D}\sigma_{ki}^2 - \frac{1}{2}\sum_i^{D}\mu_{ki}^2\Big)\Bigg] \tag{18}$$

where $\boldsymbol{\omega} = \{\omega_1,\cdots,\omega_K\}$ is an one-hot vector indicating which Gaussian component is activated.

Given that the posterior $q_\phi(\boldsymbol{z},\boldsymbol{\omega}|\boldsymbol{x})$ is assumed to be Gaussian mixture, we can calculate the second term as follows

$$\mathcal{L}_{logz\omega|x} \tag{19}$$

$$= \mathbb{E}_{\boldsymbol{z},\boldsymbol{\omega}}\Big[\log q_\phi(\boldsymbol{z},\boldsymbol{\omega}|\boldsymbol{x})\Big] \tag{20}$$

$$= \mathbb{E}_{\boldsymbol{\omega}}\Big[\int_{\boldsymbol{z}\sim q_\phi(\boldsymbol{z},\boldsymbol{\omega}|\boldsymbol{x})} q_\phi(\boldsymbol{z},\boldsymbol{\omega}|\boldsymbol{x})\log q_\phi(\boldsymbol{z},\boldsymbol{\omega}|\boldsymbol{x})d\boldsymbol{z}\Big] \tag{21}$$

$$= \mathbb{E}_{\boldsymbol{\omega}}\Bigg[\int_{\boldsymbol{z}}\frac{1}{(2\pi)^{D/2}}\frac{1}{\prod_i^D \sigma_{ki}}\boldsymbol{s}_k^{\omega_k}\exp\Big\{-\frac{1}{2}\sum_i\frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2}\Big\}\log\Big(\frac{1}{(2\pi)^{D/2}}\frac{1}{\prod_i^D \sigma_{ki}}\boldsymbol{s}_k^{\omega_k}\exp\Big\{-\frac{1}{2}\sum_i\frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2}\Big\}\Big)d\boldsymbol{z}\Bigg] \tag{22}$$

$$= \mathbb{E}_{\boldsymbol{\omega}}\Bigg[\int_{\boldsymbol{z}}\frac{1}{(2\pi)^{D/2}}\frac{1}{\prod_i^D \sigma_{ki}}\boldsymbol{s}_k^{\omega_k}\exp\Big\{-\frac{1}{2}\sum_i\frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2}\Big\}\Big(\log\frac{1}{(2\pi)^{D/2}}\frac{1}{\prod_i^D \sigma_{ki}}\boldsymbol{s}_k^{\omega_k}(-\frac{1}{2}\sum_i\frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2})\Big)d\boldsymbol{z}\Bigg] \tag{23}$$

$$= \mathbb{E}_{\boldsymbol{\omega}}\Bigg[\frac{1}{(2\pi)^{D/2}}\frac{1}{\prod_i^D \sigma_{ki}}\boldsymbol{s}_k^{\omega_k}\Big[\int_{\boldsymbol{z}}\log\Big(\frac{1}{(2\pi)^{D/2}}\frac{1}{\prod_i^D \sigma_{ki}}\boldsymbol{s}_k^{\omega_k}\Big)\exp\Big\{-\frac{1}{2}\sum_i\frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2}\Big\}d\boldsymbol{z} \tag{24}$$

$$-\int_{\boldsymbol{z}} \frac{1}{2} \sum_i \frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2} \exp\Big\{-\frac{1}{2}\sum_i \frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2}\Big\}\Big] d\boldsymbol{z}\Big] \tag{25}$$

$$= \mathbb{E}_{\boldsymbol{\omega}}\Big[\frac{1}{(2\pi)^{D/2}}\frac{1}{\prod_i^D \sigma_{ki}}\boldsymbol{s}_k^{\omega_k}\Big[\log\Big(\frac{1}{(2\pi)^{D/2}}\frac{1}{\prod_i^D \sigma_{ki}}\boldsymbol{s}_k^{\omega_k}\Big)\Big(\prod_i^D \sigma_{ki}\Big)(2\pi)^{D/2} - \frac{D}{2}\Big(\prod_i^D \sigma_{ki}\Big)(2\pi)^{D/2}\Big]\Big] \tag{26}$$

$$= \mathbb{E}_{\boldsymbol{\omega}}\Big[\boldsymbol{s}_k^{\omega_k}\Big[\log\Big(\frac{1}{(2\pi)^{D/2}}\frac{1}{\prod_i^D \sigma_{ki}}\boldsymbol{s}_k^{\omega_k}\Big) - \frac{D}{2}\Big]\Big] \tag{27}$$

$$= \mathbb{E}_{\boldsymbol{\omega}\sim\mathcal{M}(K)}\Big[\boldsymbol{s}_k^{\omega_k}\Big(-\frac{D}{2}\log(2\pi) - \frac{D}{2} - \sum_i^D \log\sigma_{ki} + \log\boldsymbol{s}_k^{\omega_k}\Big)\Big] \tag{28}$$

Below we show the key aspects of the ELBO derivation:

$$\tag{29}$$

$$-\frac{1}{2}\int_{-\infty}^{+\infty}\sum_i z_{ki}^2 \exp\Big\{\big(-\frac{1}{2}\sum_i \frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2}\big)\Big\}dz_{k1} \tag{30}$$

$$= -\frac{1}{2}\Big[\underbrace{\int_{-\infty}^{+\infty} z_{k1}^2 e^{(-\frac{1}{2}\sum_i \frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2})}dz_{k1}}_{A} + \underbrace{\int_{-\infty}^{+\infty}\sum_{i=2}^D z_{ki}^2 e^{(-\frac{1}{2}\frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2})}dz_{k1}}_{B}\Big] \tag{31}$$

Solving for A and B separately yields:

$$A = \int_{-\infty}^{+\infty} -\sigma_{k1}^2 z_{k1} de^{(-\frac{1}{2}\sum_i \frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2})} \tag{32}$$

$$= -\sigma_{k1}^2\Big[z_{k1}e^{(-\frac{1}{2}\sum_i \frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2})}\Big|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} e^{(-\frac{1}{2}\sum_i \frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2})}dz_{k1}\Big] \tag{33}$$

$$= -\sigma_{k1}^2\Big[0 - \int_{-\infty}^{+\infty} e^{(-\frac{1}{2}\sum_i \frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2})}dz_{k1}\Big] \tag{34}$$

$$= \sigma_{k1}^2\int_{-\infty}^{+\infty} e^{-\frac{1}{2}\frac{(z_{k1}-\mu_{k1})^2}{\sigma_{k1}^2}}e^{-\frac{1}{2}\sum_{i=2}^D \frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2}}dz_{k1} \tag{35}$$

$$= \sigma_{k1}^2\sigma_{k1}\sqrt{2\pi}e^{-\frac{1}{2}(\sum_{i=2}^D \frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2})} \tag{36}$$

$$= \sigma_{k_1}^3\sqrt{2\pi}e^{-\frac{1}{2}\sum_{i=2}^D \frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2}} \tag{37}$$

Similarly we can derive:

$$B = \sum_{i=2}^D z_{ki}^2\sigma_{k1}\sqrt{2\pi}e^{(-\frac{1}{2}\frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2})} \tag{38}$$

Let:

$$x_D = \int_{-\infty}^{+\infty} -\frac{1}{2}\sum_i z_{ki}^2 e^{(-\frac{1}{2}\frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2})}dz \tag{39}$$

So we can get:

$$-\frac{1}{2}[A+B] = -\frac{1}{2}\Big[\sigma_{k_1}^3\sqrt{2\pi}e^{-\frac{1}{2}\sum_{i=2}^D \frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2}} + \sum_{i=2}^D z_{ki}^2\sigma_{k1}\sqrt{2\pi}e^{(-\frac{1}{2}\frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2})}\Big] \tag{40}$$

$$= -\frac{1}{2}\Big[\sigma_{k1}^3\sqrt{2\pi}\sigma_{k2}\sigma_{k3}\cdots\sigma_{kD}(\sqrt{2\pi})^{D-1} + \sigma_{k1}\sqrt{2\pi}\sum_{i=2}^D z_{ki}^2 e^{(-\frac{1}{2}\frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2})}\Big] \tag{41}$$

$$= -\frac{1}{2}\Big[\sigma_{k1}^2\prod_{i=1}^D \sigma_{ki}(2\pi)^{\frac{D}{2}} + \sigma_{k1}\sqrt{2\pi}x_{D-1}\Big] \tag{42}$$

$$x_D = -\frac{1}{2}[\sigma_{kD}^2 \prod_{i=1}^{D} \sigma_{ki}(2\pi)^{\frac{D}{2}}] + \sigma_{kD}\sqrt{2\pi}x_{D-1} \tag{43}$$

$$x_1 = -\frac{1}{2}\sigma_1^3\sqrt{2\pi} \tag{44}$$

$$x_0 = -\frac{1}{2}(\sqrt{2\pi})^D \prod_{i=1}^{D} \sigma_{ki}(\sum_{i=1}^{D} \sigma_{ki}^2) \tag{45}$$

At this point, the equation can be converted into a problem of summing a series.

$$\int_z \frac{1}{(2\pi)^{\frac{D}{2}}} s_k^{T_k} e^{-\frac{1}{2}\sum_i \frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2}} [\log(\frac{1}{(2\pi)^{\frac{D}{2}}} s_k^{T_k}) - \frac{1}{2}\sum_{i=1}^{2} z_{ki}^2] dz \tag{46}$$

$$= \frac{1}{(2\pi)^{\frac{D}{2}}} s_k^{T_k} [\underbrace{\int_z e^{-\frac{1}{2}\sum_i \frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2}} \log(\frac{1}{(2\pi)^{\frac{D}{2}}} s_k^{T_k}) dz}_{C} + \underbrace{\int_z -\frac{1}{2}\sum_{i=1}^{D} z_{ki}^2 e^{-\frac{1}{2}\sum_i \frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2}} dz}_{D}] \tag{47}$$

$$C = \log(\frac{1}{(2\pi)^{\frac{D}{2}}} S_k^{T_k}) \int_z e^{-\frac{1}{2}\sum_i \frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2}} dz = \log(\frac{1}{(2\pi)^{\frac{D}{2}}} S_k^{T_k}) \prod_{i=1}^{D} \sigma_{ki}(2\pi)^{\frac{D}{2}} \tag{48}$$

$$D = \int_z -\frac{1}{2}\sum_{i=1}^{D} z_{ki}^2 e^{-\frac{1}{2}\sum_i \frac{(z_{ki}-\mu_{ki})^2}{\sigma_{ki}^2}} dz = -\frac{1}{2}(2\pi)^{\frac{D}{2}} \prod_{i=1}^{D} \sigma_{ki}(\sum_{i=1}^{D} \sigma_{ki}^2) \tag{49}$$

$$\tag{50}$$

Therefore, the original expression can be expressed as:

$$S_k^{Tk} \log(\frac{1}{(2\pi)^{\frac{D}{2}}} S_k^{T_k}) \prod_{i=1}^{D} \sigma_{ki} - \frac{1}{2} S_k^{T_k} \prod_{i=1}^{D} \sigma_{ki}(\sum_{i=1}^{D} \sigma_{ki}^2) \tag{51}$$

II. DETAILED DIAGNOSTIC OPINIONS FROM PROFESSIONAL OPHTHALMOLOGISTS.

Herein, we provide detailed clinical diagnostic opinions from professional ophthalmologists on the samples generated by the proposed TVPS in part.

Column 1: In the real fundus image, the optic disc color is normal, with clear boundaries. Yellow-white exudates are visible beside the macula and in the temporal side.

In the generated vessels, vascular leakage is observed in the inferotemporal side of the optic disc, with irregular vessel shapes.

In the generated fundus image, yellow-white exudations are visible in the inferotemporal side of the optic disc.

In the generated fundus image used for training, yellow-white exudations are visible in the temporal side and inferotemporal side of the optic disc.

Column 3: In the real fundus image, the optic disc color is normal, with clear boundaries. Yellow-white exudations, hemorrhages, and microaneurysms are visible beside the macula and in the peripheral retina.

In the generated vessels, vascular leakage is visible in the inferotemporal side of the optic disc. The shape of some vessels is obscured by hemorrhages, leading to an irregular course.

In the generated fundus image, yellow-white exudations are visible in the temporal side and inferotemporal side of the optic disc, along with patchy hemorrhages and microaneurysms.

In the generated fundus image used for training, yellow-white exudations and multiple hemorrhages are visible in the temporal side and inferotemporal side of the optic disc.

Column 4: In the real fundus image, the optic disc color is normal, with clear boundaries. Hemorrhages and exudations are visible in the posterior pole.

In the generated vessels, there is visible vascular deficiency in the superotemporal side of the optic disc, which appears to be due to vascular obscuration caused by hemorrhage.

In the generated fundus image, a small amount of preretinal hemorrhage and exudation is visible in the superotemporal side.

In the generated fundus image used for training, yellow-white exudations and multiple hemorrhages are visible in the temporal side and superotemporal side of the optic disc.

Column 6: In the real fundus image, the optic disc color is normal, with clear boundaries. The vessels in the superotemporal side have a tortuous shape, and flame-shaped hemorrhages are visible.

In the generated vessels, hemorrhage in the superotemporal side of the optic disc leads to obscuration and results in discontinuity of some vessels.

In the generated fundus image, hemorrhages and exudations are visible in the superotemporal side of the optic disc.

In the generated fundus image used for training, hemorrhages and exudations are visible in the superotemporal side of the optic disc, with a larger extent compared to the generated fundus images.

Column 7: In the real fundus image, the optic disc color is normal. The veins have a tortuous course, and flame-shaped hemorrhages are visible all around.

In the generated vessels, the vessels all around are obscured by hemorrhages, resulting in an incomplete course, especially severe in the superior nasal area.

In the generated fundus image, the optic disc boundaries are indistinct, with exudations visible around. The tortuous veins and a small amount of flame-shaped hemorrhages are visible all around.

In the generated fundus image used for training, flame-shaped hemorrhages are visible all around, and exudations are visible on the nasal side of the optic disc.

Column 8: In the real fundus image, the optic disc color is normal. A preretinal hemorrhage of approximately 8 PD in size is visible in the superotemporal side, and small patchy flame-shaped hemorrhages are visible in the inferotemporal side.

In the generated vessels, retinal hemorrhages in the inferotemporal side obscure the vessels, and at the same time, some vessels have an incomplete shape.

In the generated fundus image, the boundaries of the optic disc are visible. In the posterior pole, preretinal and subretinal hemorrhages of approximately 8 PD in size are visible, and exudations are visible in the periphery.

In the generated fundus image used for training, the optic disc color is normal. A preretinal hemorrhage of approximately 8 PD in size is visible in the superotemporal side, and small patchy exudations are visible in the inferotemporal side.

Column 10: In the real fundus image, the optic disc color is normal, with clear boundaries. A scar-like lesion is visible in the macula, accompanied by a small amount of hemorrhage in the center.

In the generated vessels, vascular deficiency is visible in the inferotemporal side of the optic disc, seemingly due to vascular obscuration caused by hemorrhage.

In the generated fundus image, a scar is visible in the macular area, and hemorrhage is visible in the inferotemporal side of the scar.

In the generated fundus image used for training, a scar is visible in the macular area, and hemorrhage is visible in the inferotemporal side of the scar.

Column 11: In the real fundus image, the optic disc color is normal, with clear boundaries. Large areas of exudation are visible in the temporal side of the macula, along with preretinal and subretinal hemorrhages.

In the generated vessels, vascular leakage is visible in the superotemporal side of the optic disc. The course of some vessels is obscured by hemorrhages, leading to an irregular shape, or in other words, there is partial occlusion of the vessels in the superotemporal side.

In the generated fundus image, yellow-white exudations are visible in the temporal side of the optic disc, and subretinal hemorrhage is visible in the temporal side of the macula.

In the generated fundus image used for training, multiple yellow-white exudations are visible in the temporal side and inferotemporal side of the optic disc, and preretinal and subretinal hemorrhages are visible in the temporal side of the macula.

Column 12: In the real fundus image, the optic disc color is normal. Hemorrhages are visible in the macular area, and

exudations are visible around the macula.

In the generated vessels, the shape of the vessels in the temporal side of the optic disc is irregular.

In the generated fundus image, the boundaries of the optic disc are visible. A scar of approximately 2/3*2/3 PD in size is visible in the macular area, accompanied by surrounding hemorrhages.

In the generated fundus image used for training, the optic disc color is normal. Hemorrhages are visible in the macular area, accompanied by surrounding exudations, and a scar is visible on the nasal side of the macula.

III. DETAILS OF THE NETWORK ARCHITECTURE

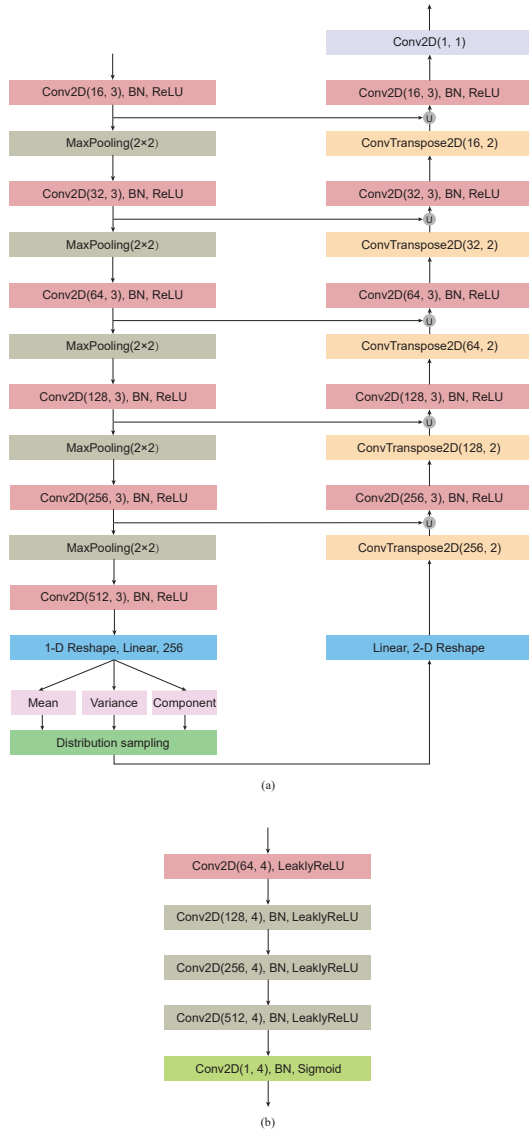Here, we present the specific architectures of the encoder-decoder and discriminator utilized in TVPS.

Fig. S1. Network architectures of the *neural encoder F_ϕ and decoder G_θ* (a) and *adversarial discriminator D_ψ* (b) in the proposed TVPS. Ⓤ stands for feature concatenation.