1. The Main Contributions of This Paper

- This paper presents a novel linear prediction-based part-defined auto-encoder (PAE) network.
- The parallel Network is used to estimate the modification factor of AR-wiener filter mask.
- The PESQ and STOI results of the LP-based PAE are better than baseline methods at lower signal noise ratio (SNR) levels.

2. PAE-based Speech Enhancement

3. The Loss Function of The PAE

- **PAE based on the AR filter**
  
  The Decoder as synthesizer is based on the AR Wiener filter
  
  \[ H_{\text{AR}}(k) = \frac{A(k)}{e^t A(t) e^t} \]  

  - **The loss function based on the Eq. (1)**
  - \[ \mathcal{L} = \left[ \sum_{t=1}^{T} \sum_{n=1}^{N} R_{\text{PN}}(X_t, Y_t, W_n, k) - \sum_{t=1}^{T} \sum_{n=1}^{N} R_{\text{PN}}(X_t, Y_t, W_n, k) \right] \]

- **Modified AR-Wiener filter with the residual**

  Linear prediction residual with AR model an infinite impulse response (IIR) filter
  
  \[ g_x(n) = x(n) - \sum_{l=1}^{L} \alpha_l x(n-l) + h_n(p) \]

  In power spectral density (PSD):
  
  \[ X(k) = \frac{R(k)}{\mathcal{H}_{\text{Wiener}}(k)} \]

  - The modified AR-Wiener filter and its approach estimation
  
  \[ H_{\text{AR, Wiener}}(k) = H_{\text{AR}}(k) H_{\text{Wiener}}(k) \]  

4. Performance Evaluation

- **Experimental setup**
  
<table>
<thead>
<tr>
<th>Input</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speech</td>
<td>Enhanced Speech</td>
</tr>
</tbody>
</table>

- **4. Performance Evaluation**

  - **Performance Evaluation Results**
    
    | Noise Type | Average | Average | Average | Average |
    |-----------|--------|--------|--------|--------|
    | Babble    | 1.754  | 2.124  | 2.057  | 2.032  |
    | Factory   | 1.754  | 2.124  | 2.057  | 2.032  |
    | Office    | 1.680  | 2.167  | 1.975  | 2.149  |
    | Street    | 1.468  | 2.453  | 2.184  | 2.644  |

- **5. Conclusions and Future Work**

  - **Conclusions**
    
    - The PAE can estimate the wiener mask based on the AR model constraint.
    - The proposed neural network concentrate more spectrum structure based on the AR model.

  - **Future Work**
    
    - Learning the perception of the AR model and the mask design based on PAE
    - The structure design or the voiced speech model for better residual estimation.