

文章编号: 2095-2163(2022)03-0128-06

中图分类号: TP183

文献标志码: A

深度果蝇神经网络及其实时能见度预测

肖应慧¹, 张著洪²

(1 贵州大学 大数据与信息工程学院; 2 贵州大学 贵州省系统优化与科学计算特色重点实验室, 贵阳 550025)

摘要: 能见度预测属于时间序列预测领域的问题, 如何实现对时序数据的处理以及捕获数据中的时间长期依赖关系, 是当前研究人员的主要关注焦点。针对能见度预测模型构建难的问题, 依据果蝇视觉系统在图像感知、学习和信息反馈过程中的生物学原理, 提出一种基于视频监控图像的深度果蝇神经网络。首先, 通过设计前馈果蝇视觉神经网络, 提取图像的能见度特征, 进而将特征信息送入多层感知器, 获得关于未知权值和阈值的能见度与视频帧的映射关系; 其次, 借助梯度下降法获得在线预测能见度的深度神经网络。实验结果表明, 该神经网络能实时、有效、准确地预测雾天环境下的能见度, 预测效果稳定且精度高, 具有较好的应用潜力。

关键词: 果蝇视觉系统; 神经网络; 能见度预测; 梯度下降; 时间序列

Deep fly neural network and its real-time visibility prediction

XIAO Yinghui¹, ZHANG Zhuhong²

(1 College of Big Data and Information Engineering, Guizhou University, Guiyang 550025, China; 2 Guizhou Provincial Characteristic Key Laboratory of System Optimization & Scientific Computation, Guizhou University, Guiyang 550025, China)

[Abstract] Visibility prediction is a problem in the field of time series prediction, and how to implement the processing of time-series data and capture the temporal long-term dependence in the data is the main focus of current researchers. In order to solve the difficulty of constructing visibility prediction models, a deep fly neural network, based on video surveillance images is proposed in terms the biological principles of the fly visual system such as image perception, learning and information feedback. In the design of the model, a feedforward fly visual neural network is designed to extract the image-based visibility feature, and then the acquired feature information is fed into a multi-layer perceptron to obtain the mapping relationship between visibility and video frames with regard to unknown weights and thresholds. These, together with the gradient descent method, derive the deep neural network for real-time visibility prediction. The comparative experimental results show that the neural network can accurately predict the visibility of the foggy environment, and has great application potential.

[Key words] fly visual system; neural network; visibility prediction; gradient descent; time series

0 引言

大气能见度被定义为视力正常的人在白天以靠近地平线的天空为背景, 观测黑色物体时易识别和辨认的最长距离^[1]。能见度的变化取决于大气的透明度, 不良天气现象会导致大气浑浊, 使大气透明度降低, 特别在诸如雾、雨或空气污染等恶劣天气条件下, 能见度较低。近年来, 由雾造成低能见度而导致交通事故频繁发生, 因此准确、高效地对雾天能见度进行实时预测尤为重要。目前, 能见度预测的研究方法主要有应用统计分析法和深度学习法。统计分析法通常有物理建模法和非参数建模法, 其经由图像构建物理或数据驱动的概率模型, 估计不同雾范围下的大气能见度。物理模型法是基于视觉分析

的光反射、折射等物理现象, 以及基于 Koschmieder 定理, 建立光传播物理模型, 进而利用消光系数实时估计能见度, 能见度估计的准确率高很大程度上依赖于原始图像和无雾参考图像的相似程度^[2-3]; 非参数建模法考虑了低能见度情况下图像退化对图像分类的影响。研究表明, 能见度可由雾的空间特征进行表征。例如, Nnolim^[4] 将全局对比度算子与基于分阶的多尺度滤波器结合, 将熵和标准偏差融合, 对图像进行去雾和图像增强, 获得多尺度融合去雾算法, 但该算法虽能提升对雾天能见度估计的精度, 但不能实时预测能见度。深度学习因能对雾天能见度时序数据在线处理, 能捕获数据的中长期依赖关系, 已受到广泛关注, 并在诸如气象预报等能见度预测中已得到初步应用^[5-6]。例如, Selvaraj 等^[7]

基金项目: 国家自然科学基金(62063002, 61563009)。

作者简介: 肖应慧(1995-), 女, 硕士研究生, 主要研究方向: 智能优化算法; 张著洪(1966-), 男, 博士, 教授, 博士生导师, 主要研究方向: 数据科学与计算智能、深度学习。

通讯作者: 张著洪 Email: zhzhong@gzu.edu.cn

收稿日期: 2021-10-18

以最高温度、最低温度、最高相对湿度、最小相对湿度和风速增强长为输入量,利用长短时记忆神经网络预测降雨量; Akmaljon 等^[8]借助快速傅里叶变换 (FFT) 滤波器去除输入图像的低层次特征,并对每个输入应用频谱滤波器来提取低对比度区域,进而利用三流集成卷积神经网络获得能见度深度学习模型。

综上,现有能见度预测方法主要集中于如何从输入图像中获取有价值的能见度特征。相对而言,深度学习法比统计分析法在能见度预测中更有应用前景,但因网络结构复杂、计算效率低、易于出现过饱和或欠拟合现象,使得其应用于能见度的研究有待深入开展。为此,本文在已获针对碰撞检测的果蝇前馈神经网络基础上,从果蝇视觉感知和学习的

行为特性出发,尝试性提出一种深度果蝇神经网络 (Deep Fly Neural Network, DFNN),并应用于实时能见度预测。

1 果蝇神经信息处理

果蝇有两只复眼,每只复眼由大约 3 500 个小眼组成,且每个小眼包含 8 个光感受器。光信息通过复眼进入果蝇的视觉系统,实现光传递。果蝇视觉系统具有独特和简单的眼结构,对外界环境信息的感知能力强,其神经系统依次经由光感受器 (Photoreceptor)、视网膜 (Retina)、薄膜层 (Lamina)、髓质 (Medulla)、小叶 (Lobula) 及大脑神经信息处理层,执行视觉信息的处理和传递,各层之间分工明确又联系紧密,如图 1 所示。

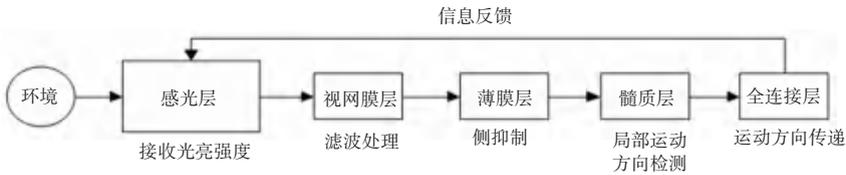


图 1 果蝇视觉信息处理的流程图

Fig. 1 Flow chart of visual information processing in fly

光感受器层 (Photoreceptor) 主要由感光细胞组成,负责接收外来的光强度信息;视网膜层 (Retina) 接收光感受器层输出的信号,负责消除视觉噪声信号;薄膜层又称为视叶神经层 (Lamina), 包含墨盒 cartridge (cart) 和开关 on-off (oo) 这两种类型节点,其中 cart 节点接收来自视网膜上对应节点的邻域节点的输入,并通过卷积滤波产生输出信号,而 oo 节点负责接收对应 cart 节点邻域内节点的输出信号,并经由超极化和去极化处理,产生行为输出量;髓质层包含柱状神经元和切向细胞,由主通道子单元的两类结点 (m_1 、 m_2) 组成,其中 m_1 节点接收薄膜层中对应的 oo 节点的输出,并输出检测节点处的局部方向; m_2 节点接收 m_1 节点邻域处的局部方向量,经由投影产生在水平、竖直方向的方向量;小叶层汇集髓质层中所有 m_2 节点的水平、竖直方向量,并将信息整合后输入果蝇大脑神经系统。最后,大脑神经通过接收小叶层的信息以及神经元相互作用机制,获得视觉信息刺激后的反馈信号,并将误差信号反馈入光感受器层。

2 深度果蝇神经网络

基于果蝇视觉系统的信息处理机制,针对视觉场景下的碰撞检测与预警问题,获得一种人工果蝇

视觉神经网络 (Artificial Fly Visual Neural Network, AFVNN), 其由 4 个神经层构成^[9]。在此,对该神经网络作适当改进,得到用于实时能见度预测的改进型深度果蝇神经网络 (Deep Fly Neural Network, DFNN)。DFNN 由提取图像特征的视觉神经网络和多层感知器 (Multiple Layer Preceptor, MLP) 构成前向传播神经网络,经由梯度下降更新网络中卷积参数、权值和阈值,实现误差反向传播。改进型果蝇深度神经网络结构如图 2 所示。

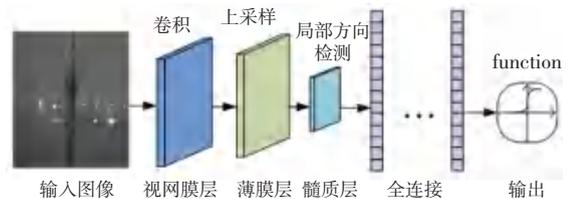


图 2 果蝇深度前馈神经网络的示意图

Fig. 2 Schematic diagram of the Fly deep feedforward neural network

给定由时间序列灰度图与标签构成的训练集 $\{(A_t, Y_t)\}_{t \geq 1}$, A_t 为第 t 时刻的 $M \times N$ 灰度图, Y_t 为 n_0 维的标签或观测值向量。经由图 2 获知,DFNN 接收连续的视频图像帧,并依次经由视网膜、薄膜层、髓质层这 3 个视觉神经层提取输入图像的特征信息,进而经由 MLP 产生神经网络的输出信号。

2.1 前馈传播

2.1.1 Retina 层

该层由 $M \times N$ 个 Retina 节点形成的 M 行和 N 列矩阵表示,在第 t 时刻接收灰度图像 A_t ,在此 $3 \mid (M - 2)$, $3 \mid (N - 2)$ 。Retina 节点 (i, j) 在第 t 时刻接收输入图像中对应像素点处的灰度值 $L_{ij}(t)$ (光亮强度),并经由式(1) 输出光亮强度偏差信号:

$$G_{ij}(t) = \frac{1}{2}(L_{ij}(t) - g(L_{ij}(t))), 1 \leq i \leq M, 1 \leq j \leq N \quad (1)$$

其中, $g(\cdot)$ 表示时滞函数。

2.1.2 Lamina 层

此层由 cartridge 和 on-off 两个子层构成。cartridge 层由 $(M - 2) \times (N - 2)$ 个 cart 节点构成,依据窗口大小为 3×3 、步长为 1 的滑动方式,每个节点接收对应的 3×3 个 Retina 节点的输出,并经由高斯滤波器产生该节点的输出膜电位,即式(2):

$$X_{rs}(t) = f\left(\sum_{u,v=0}^2 w_{uv} G_{r+u,s+v}(t)\right) \quad (2)$$

其中, $1 \leq r \leq M - 2$; $1 \leq s \leq N - 2$; w_{uv} 是待定的 3×3 权重矩阵 \mathbf{W}_1 中位置 (u, v) 处的元素; $f(\cdot)$ 是非线性激活函数 $Relu$ 。类似地, on-off 层由 $(M - 2) \times (N - 2)$ 个 oo 节点构成,其依据扩边方式和窗口大小为 3×3 、步长为 1 的滑动方式,每个节点接收对应的 3×3 个 cart 节点的输出,并经由如下侧抑制机制,输出其膜电位,式(3):

$$x_{rs}(t) = E_{rs}(t) - \frac{1}{8} \sum_{0 \leq u,v \leq 2} g(E_{r+u-1,s+v-1}(t)), u,v \neq 1 \quad (3)$$

在此, $E_{rs}(t)$ 是 cartridge 层中节点 (r, s) 输出的膜电位的偏差量,即式(4):

$$E_{rs}(t) = X_{rs}(t) - g(X_{rs}(t)) \quad (4)$$

2.1.3 Medulla 层

此层由 $p \times q$ 个 m 节点按矩阵排列表示,其中 $p = (M - 2)/3$, $q = (N - 2)/3$ 。首先,将 Lamina 层中按矩阵排列的 $(M - 2) \times (N - 2)$ 个 oo 节点划分为 $p \times q$ 个子块。基于对称 EMD 运动方向检测器和 3×3 权值矩阵 \mathbf{W} , m 节点 (i, j) 经由式(5) 输出膜电位:

$$m_{ij}(t) = \sum_{0 \leq u,v \leq 2} \mathbf{W}(x_{3i+1,3j+1}g(x_{3i+u,3j+v}) - x_{3i+u,3j+v}g(x_{3i+1,3j+1})), 1 \leq i \leq p, 1 \leq j \leq q, u,v \neq 1 \quad (5)$$

其中, \mathbf{W} 是一个大小为 3×3 的权重矩阵:

$$\mathbf{W} = \begin{bmatrix} \frac{\alpha\sqrt{2}}{2} & -1 & -\sqrt{2}\bar{0} \\ \frac{\alpha}{2} & 1 & -1 \\ \frac{\alpha}{2} & 1 & 0 \\ \frac{\alpha\sqrt{2}}{2} & 1 & 0 \end{bmatrix}$$

2.1.4 全连接层

Medulla 层输出的特征图矩阵通过 Vec 算子按列展开后形成一个列向量 $\vec{p} = (p_1, p_2, \dots, p_n)^T$, 并作为 MLP 的输入;全连接层中第 k 个隐含层的第 i 个神经节点的输出为式(6):

$$a_i^{(k)} = f(z_i^{(k)}) = f\left(\sum_{m=1}^{n_{k-1}} \mathbf{w}_{mi}^{(k-1)} a_m^{(k-1)} + b_i^{(k)}\right) \quad (6)$$

$1 \leq i \leq n_k, 1 \leq k \leq K$

其中, $z_i^{(k)}$ 为第 k 层中第 i 个节点的输入, n_k 为第 k 隐含层的节点数。输出层的输出为式(7):

$$\hat{Y}_j = O(z_j^{(o)}), 1 \leq j \leq n_o \quad (7)$$

其中, $O(x) = 1.52x + 2.3$, n_o 为输出层的节点数。

在第 t 时刻,前馈神经网络的输出误差函数由式(8)表示:

$$E_t(\mathbf{W}_1, \mathbf{W}_2, \mathbf{b}) = \frac{1}{l} \sum_{j=t-l+1}^t \frac{1}{n_o} \sum_{j=1}^{n_o} (\hat{Y}_j - Y_j)^2 \quad (8)$$

其中, \mathbf{W}_1 为 Lamina 层中权重矩阵; \mathbf{W}_2 与 \mathbf{b} 分别是 MLP 中连接权值和阈值构成的矩阵; $Y_j(t)$ 为标签值; l 为一个训练周期的批量大小。

2.2 反向传播

由式(8)可知, FVNN 的反向传播是将前馈传播的误差量反向传递给各神经层,并更新全连接层中的权值、阈值和 Lamina 层的权值。Lamina 层的权值由式(9)更新:

$$\mathbf{W}_1^{(t+1)} = \mathbf{W}_1^{(t)} - \eta_1 \frac{\partial E^{(t)}}{\partial \mathbf{W}_1} \quad (9)$$

其中, η_1 是给定的学习率。全连接层的权值与阈值经由式(10)~(11)更新:

$$\mathbf{W}_2^{(t+1)} = \mathbf{W}_2^{(t)} - \eta_2(t) \frac{\partial E^{(t)}}{\partial \mathbf{W}_2} + \alpha \frac{\partial E^{(t-1)}}{\partial \mathbf{W}_2} \quad (10)$$

$$\mathbf{b}^{(t+1)} = \mathbf{b}^{(t)} - \eta_2(t) \frac{\partial E^{(t)}}{\partial \mathbf{b}} + \alpha \frac{\partial E^{(t-1)}}{\partial \mathbf{b}} \quad (11)$$

其中, $\alpha \in (0, 1)$, 学习率 $\eta_2(t)$ 依据前馈误差作自适应调节,即式(12):

$$\eta_2(t) = \begin{cases} k_{inc} \eta_2(t-1), E^{(t)} < E^{(t-1)} \\ k_{dec} \eta_2(t-1), E^{(t)} \geq E^{(t-1)} \end{cases} \quad (12)$$

其中, k_{inc} 和 k_{dec} 为增量和减量因子。

2.3 DFNN 的算法描述

结合以上各层的设计, DFNN 的算法描述如下:

- (1) 参数设置: 灰度图大小 $M \times N$; 学习率 η_1 ; 参数 k_{inc} 和 k_{dec} ;
- (2) 置 $t \leftarrow m, m > l, l = m$, 随机初始化 $\mathbf{W}_1, \mathbf{W}_2$ 及 \mathbf{b} ;
- (3) 输入第 t 时刻的样例 (A_t, Y_t) ;
- (4) 依据式(8)计算前馈误差值;

(5) 经由式(9)~式(11)更新网络中的权值和阈值,并输出第 $t+1$ 时刻的能见度预测值;

(6) $t \leftarrow t + 1$;若不满足终止条件,则返回(3);否则,则结束。

DFNN 的算法计算复杂度主要由步 4~5 确定。在一个迭代周期内, Retina 和 Lamina 层分别进行 $2LMN$ 和 $20L(MN - 2M - 2N + 4)$ 次运算; Medulla 层共进行 $L(13MN - 26M - 26N + 52)/3$ 次运算;全连接层共运算 $2(MN - 2M - 2N + 22)/9$ 次。另一方面,在反向传播中,全连接层共进行 $(3MN - 6M - 6N + 66)/9$ 次信息传播操作; Medulla 层共进行 $(10MN - 20M - 20N + 40)/9$ 次乘法运算; Lamina 层共进行 $9 + (82MN - 164M - 164N + 409)/9$ 次运算。因此,DFNN 在一个时间周期内,前向传播的计算总次数为

$$\frac{1}{9}((234MN - 366M - 366N)l + 876) \quad (13)$$

反向传播的计算总次数为

$$\frac{1}{9}((95MN - 190M - 190N) + 515) \quad (14)$$

由此可知,DFNN 的计算复杂度主要由输入图像分辨率 M 、 N 和 l 确定;因此,只要适当设置 M 、 N 和 l 的取值,DFNN 的运行效率能得到保障。

3 数值实验

在 Windows10 (CPU/i3 4170 3.70 GHz, RAM/4 GB)/Python 3.6 环境下展开数值实验。为验证神经网络 DFNN 应用于能见度预测问题的有效性,选择

经典的卷积神经网络 CNN、多层感知器 MLP 以及改进型的 VisNet 参与比较。各算法的最大迭代次数均为 1 858。经参数调试,DFNN 的参数设置为 $k_{inc} = 2$, $k_{dec} = 0.5$, $l = m = 4$, $\eta_1 = 0.01$, $\eta_2 = 0.1$; CNN 的卷积层和池化层的层数均为 5,隐含层数 5; CNN 与 MLP 网络的激活函数为 *Relu* 和一元线性函数 $O(\cdot)$, 学习率为 0.01; VisNet 的参数设置源于相应的文献。

3.1 模型指标

为评估模型的性能,各模型的输出和预测值时间序列均利用已有的均方根误差 (Root Mean Square Error, *RMSE*)、平均绝对误差 (Mean Absolute Error, *MAE*)、平均绝对百分比误差 (Mean Absolute Percentage Error, *MAPE*)、决定系数 (Coefficient of Determination, R^2) 以及准确率 (Accuracy, *Acc*) 进行模型评价。*RMSE* 作为一种绝对误差度量指标,刻画模型的离散程度,该值越接近 0,则模型的稳定性越好; *MAE* 度量模型的预测能力,其值越小,则模型的预测精度越高; *MAPE* 是基于相对比误差的精度度量,刻画预测结果与观测值的平均偏离程度; R^2 表征预测值与观测值之间的相关性,其越接近 1,则预测值与观测值的时间序列的关联度越高; *Acc* 的值越接近 1,则模型的输出值或预测值与观测值的误差越小。

3.2 测试事例及分析

基于 2020 年全国研究生数学建模竞赛 E 题提供的视频帧和能见度标签数据,选取分辨率为 320×320 的 1 862 幅灰度图及对应的观测能见度值构成样本集 Σ , 每张图像内包含目标物灯塔和天空信息,部分灰度图如图 3 所示。

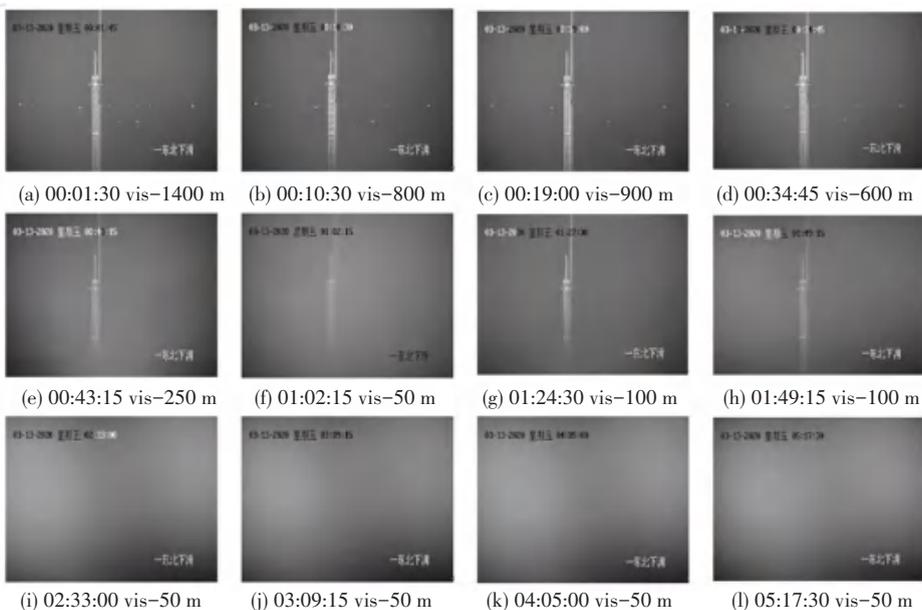


图 3 不同能见度下样板图像示例

Fig. 3 Example of sample images at different visibility levels

由图3的视觉效果可知,随着时间的推移,雾的浓度越来越大,目标物体灯塔的可见性越来越差,且逐渐模糊,即图像的能见度距离逐渐变低。该视频在较长时间段内,目标物体的能见度都很低。

3.3 实验结果与分析

将样本集 Σ 中的样例作为 DFNN 及参与比较模型的输入,获得各模型输出神经元输出的能见度值和下一时刻的预测值,依据评价指标获得各自的评价指标值,见表1。另外,各模型作用于 Σ 中时间

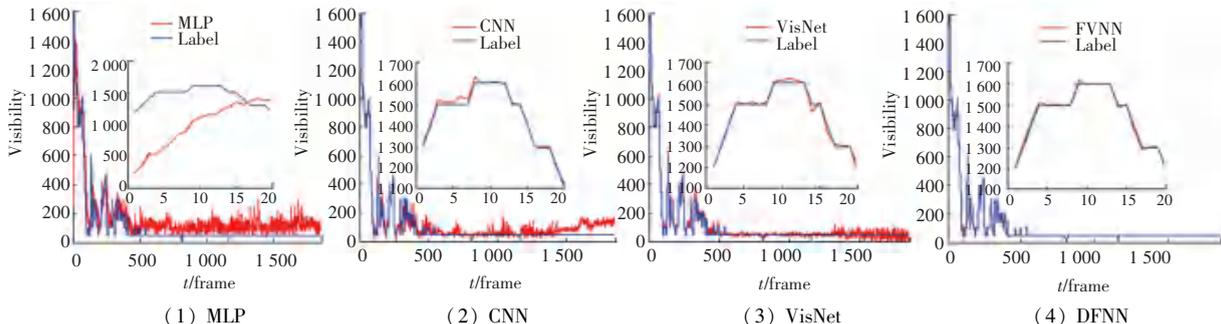


图4 不同能见度下模型的能见度预测值与观测值对比

Fig. 4 Modeled visibility predictions versus observed values at different visibility levels

由表1可知,DFNN获得的均方根误差值 $RMSE$ 最小且与其它模型获得的值偏差较大,说明DFNN获得的能见度预测值时间序列与能见度观测时间序列的离散程度较低,因而其预测雾天环境下能见度的稳定性较好;MLP获得的 $RMSE$ 值最大,说明将输入灰度图的灰度值直接用于训练神经网络模型会导致MLP的预测值与观测值的离散程度高,模型预测的稳定性较差;CNN和VisNet获得的 $RMSE$ 相对于MLP的值偏小,说明将输入灰度图作特征提取有助于改善模型预测效果的稳定性。由各模型获得的 MAE 指标值可知,DFNN的能见度预测效果最好,VisNet次之,而MLP的预测效果最差。由 $MAPE$ 指标值获知,DFNN获得的 $MAPE$ 值接近0,表明其在各个时刻获得的能见度预测值较为逼近观测值,但其它模型却相反,即在特定的时刻或时间段出现预测误差偏大。另外,DFNN和VisNet得到的 R^2 值已接近1,说明这两种模型获得各时刻的能见度预测值与观测值高度相关,因此模型的拟合能力强;虽然MLP和CNN得到的能见度预测值与观测值也具有一定的相关性,但模型的拟合能力相对较差。最后,指标 Acc 的值表明,DFNN的模型预测精度高,VisNet次之,而MLP最差。概括起来,DFNN具有最好的实时能见度预测效果,其预测精度高且稳定,VisNet次之,且MLP对实时预测能见度的能力偏弱。

序列样例,获得的能见度预测曲线与观测值曲线比较如图4所示。

表1 算法的误差性能比较

Model	$RMSE$	MAE	$MAPE$	R^2	$Acc / \%$
MLP	2.103 6	1.476 9	1.299 9	0.909 5	60.23
CNN	0.841 7	0.574 3	0.934 1	0.851 2	81.81
VisNet	0.364 4	0.232 2	0.340 9	0.992 9	93.96
DFNN	0.052 0	0.014	0.000 3	0.999 7	98.51

由图4可知,因MLP未涉及图像的特征提取,所以其解决能见度时间序列预测问题的能力较弱,且预测效果出现明显不稳定的现象;CNN的能见度预测效果比MLP的要好,但预测效果的波动性较为明显,且在预测后期,预测曲线与观测值曲线存在一定的偏离;VisNet的预测效果比CNN要好,且预测值与观测值的偏差较小,但预测曲线在后期出现波动现象;DFNN明显优于参与比较的模型,能快速地在预测初期使预测值与观测值的偏差变小,所得预测曲线逼近观测值曲线的程度较高,且预测效果未出现波动现象,因此其较适合基于图像的预测问题。

4 结束语

能见度估计一直是交通出行与管理关注的重要话题,也是计算机视觉领域的重要课题。本文在分析果蝇视觉信息处理与传递的生物学原理基础上,基于果蝇的视觉信息处理机制和学习特性,以及梯度下降算法,获得一种深度果蝇神经网络模型。计算复杂度分析表明,该模型的计算代价主要由其输入分辨率确定。实验结果表明,此模型应用于时间序列能见度预测问题,具有较好的预测能力和抗干扰能力,预测效果稳定,实时能见度预测精度高。此外,该模型虽能满足实时性要求,但环境的自适应能力有待进一步开展研究。